Structural and biophysical Analysis of plasmodial and bacterial multimeric Pdx1 - Pdx2 Proteins and Complexes

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften (*Dr. rer. nat.*)

an der Fakultät für Mathematik, Informatik und Naturwissenschaften

der Universität Hamburg

Fachbereich Chemie

vorgelegt von

Najeeb Ullah

Hamburg, March 2021

Certification

The present work was carried out in the period from May 2017 to November 2020 in the group of Prof. Christian Betzel in the Laboratory for Structural Biology of Infection and Inflammation (DESY), Institute of Biochemistry and Molecular Biology, Department of Chemistry at the University of Hamburg.

Gutachter-1: Prof. Dr. Christian Betzel

Gutachter-2: Prof. Dr. Andrew Torda

Tag der Disputation: 16.04.2021

Dedicated to my wife and children

Publications

1. **Najeeb Ullah,** Hina Andaleeb, Celestin N. Mudogo, Sven Falke, Christian Betzel & Carsten Wrenger, "Solution Structures and Dynamic Assembly of the 24-Meric Plasmodial Pdx1–Pdx2 Complex" Int. J. Mol. Sci. 21 (17), 5971, (2020).

2. Hina Andaleeb, **Najeeb Ullah**, Sven Falke, Markus Perbandt, Hévila Brognaro & Christian Betzel, "High-resolution crystal structure and biochemical characterization of a GH11 endoxylanase from *Nectria haematococca*". Sci. Rep. 10, 15658, (2020).

Table of Contents

Abbreviations

AMR	Antimicrobial resistance
ASU	Asymmetric unit
AUC	Analytical ultra-centrifugation
BZU	Bahauddin zakariya university
Cα	Alpha carbon
CD	Circular dichroism
CSS	Complex formation significance score
CSSB	Center for structural systems biology
DESY	Deutsches electronen-synchrotron
DLS	Dynamic light scattering
Dmax	Maximum diameter
DSSP	Define the secondary structure of proteins
DTT	Dithiothreitol
DXP	Deoxy-D-xylulose 5-phosphate
EDTA	Ethylenediamine-tetra-acetic acid
EM	Electron microscopy
EMBL	European molecular biology laboratory
G3P	Glyceraldehyde 3-phosphate
GATase	Glutamine amidotransferase
GTS	Global technical strategy
GUI	Graphical-user-interface
HEC	Higher education commission
I(s)	Scattering intensity
IPTG	Isopropyl β-d-1-thiogalactopyranoside
LiCl	Lithium chloride
MM	Molecular masses
MP	Mass photometry
MR	Molecular replacement
MRSA	Methicillin-resistant *Staphylococcus aureus*
MTG	Monothioglycerol
MWCO	Molecular weight cut-off
Ni-NTA	Nickel nitrilotriacetic acid
NTA	Nanoparticle tracking analysis
PAGE	Polyacrylamide gel electrophoresis
PDB	Protein data bank

Pdx	Pyridoxal
Pdx1	Pyridoxal phosphate synthase domain
Pdx2	Pyridoxal phosphate glutaminase domain
Pr	Distance distribution function
PLP	Pyridoxal phosphate
PMSF	Phenylmethylsulfonylfluoride
Pv	*Plasmodium vivax*
R5P	Ribose 5-phosphate
R_G	Radius of gyration
R_H	Hydrodynamic radius
RMSD	Root mean square deviation
Sa	*Staphylococcus aureus*
SAXS	Small-angle X-ray scattering
SDS	Sodium dodecyl sulfate
SEC	Size-exclusion chromatography
SOPMA	Self-optimized prediction method with alignment
SOR1	Singlet oxygen resistance 1 gene
TEV	Tobacco etch virus
TIM	Triosephosphate Isomerase
UV	Ultraviolet light
XDS	X-rays detector software
ρ	Shape factor

1.0. Abstract

The main pre-requisite of the inhibitor-based selection approaches to cure diseases and infections is based on the evidence that the pathogen drug target must not be present in the host and/or the potential inhibitor must be uncommon in the metabolic pathways of the host to avoid detrimental effects with the host proteins. In ideal cases, new drugs should be developed based on specifically targeting the parasite with minimal or no toxicity to humans. In this context, the essential presence of vitamin B_6 biosynthesis enzymes in the metabolic pathways of a majority of pathogenic organisms, and their absence in humans recognized the pyridoxal phosphate (PLP) biosynthesis enzyme complex as a potential drug target. In this context, my thesis research activities were focused on the structure-function-analysis of the PLP synthase complex. The main structural unit and the catalytic center of the PLP synthase complex are constituted by the Pdx1 subunit (PLP synthase domain) forming a dodecamer and Pdx2 (glutaminase domain), which binds independently to Pdx1, establishing a 24meric Pdx1-Pdx2 (Pdx) complex.

The essential importance of the PLP synthase complex is based on its structural elements controlling the entry of the substrate to the active site and also the Pdx1-Pdx2 interface regions. Utilizing the PLP synthase enzyme complex as a potential drug target for drug discovery approaches requires information about these important structural elements and/or about the Pdx1-Pdx2 interface regions. The PLP synthase complex enzyme shares its substrates (glutamine, and sugars i.e. ribose 5-phosphate and glyceraldehyde 3-phosphate) with essential host enzymes like glutamine amido-transferase, ribose 5-phosphate isomerase, and glyceraldehyde 3-phosphate dehydrogenase, involved in different metabolic pathways. Therefore, instead of substrate-based inhibitor design targeting the PLP synthase enzymes, Pdx1-Pdx2 interface-based inhibitors, is a wise direction.

Considering the potential of PLP synthase enzymes for novel drug development studies, the first part of my thesis is focused on analysing the *Plasmodium vivax* PLP synthesis proteins (Pdx1, Pdx2, and the Pdx complex) via biophysical characterization and structural investigations. As plasmodial species, protozoan pathogens, are causing severe and fatal malaria diseases and have developed an effective resistance mechanism against most of the currently available antimalarial drugs. This instigates an urgent need

to identify new antimalarial drug targets. To characterize the PLP enzymes and their complex in terms of initial drug discovery investigations, the *Plasmodium vivax* Pdx1, Pdx2, and the Pdx1–Pdx2 (Pdx) complex were investigated by utilizing complementary biophysical and bioanalytical techniques, such as dynamic light scattering (DLS), mass photometry (MP), small-angle X-ray scattering (SAXS), and electron microscopy (EM). In addition, X-ray crystallography structural studies were also carried out for the Pdx1 and Pdx complex applying single-crystal diffraction and utilizing synchrotron radiation at DESY.

The biophysical investigations revealed a dodecameric Pdx1 and a monodisperse Pdx complex. However, Pdx2 was identified in monomeric and in different oligomeric states in solution. Interestingly, mixing oligomeric and polydisperse Pdx2 with dodecameric monodisperse Pdx1, investigated by time-resolved DLS and by nanoparticle tracking analysis (NTA), resulted in the gradual disappearance of oligomeric Pdx2 and the formation of a monodisperse Pdx complex with completion in hours. The data revealed the importance of the time-resolved investigation in examining the dynamics of a complex reaction process. Subsequently, the electron microscopy studies revealed the formation of an unsaturated Pdx1-Pdx2 complex with saturation observed with the oligomeric Pdx2. In addition, the crystal structure investigations also revealed a dodecameric PvPdx1, indicating that the dodecameric form of Pdx1 is explicitly a particular feature of plasmodial species.

With no doubt, the broad-spectrum antibiotic Penicillin discovery by Alexander Fleming, is one of the utmost achievements in medicine’s history saving countless lives throughout the world over decades. But, the natural resistance development in microbes due to the overuse of the antibiotics, and long-term exposure to these antibiotics result in losing effectiveness very rapidly. Because of the very fast adaptation ability, Methicillin-resistant *Staphylococcus aureus* (MRSA) which often appeared in hospital-acquired infections, has become the most life-threatening microbe. With only a few available therapeutic possibilities, its resistance to the currently available antibiotics has been increased. Therefore, concerned institutions and also more than 20 leading biopharmaceutical companies announced the “**AMR Action Fund**” in July 2020 with an aim of bringing two-four new antibiotics or alternative therapeutics into the market by the year 2030 [1]. Concerning the alarming side effects of antibiotics and issues like

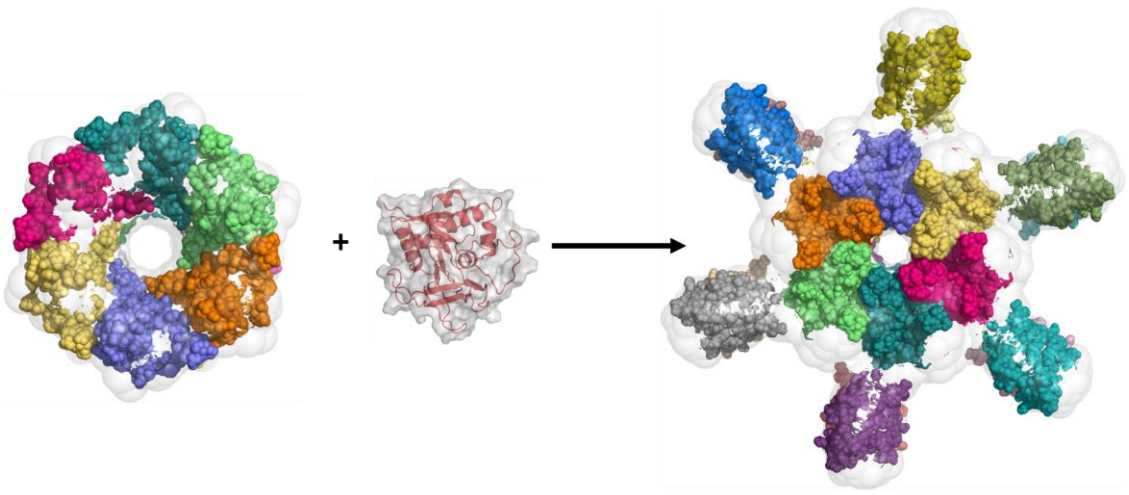
multidrug resistance, the novel approach of targeting the metabolic pathways of the microbes, particularly those absent in humans, like vitamin B_6 synthesizing Pdx complex enzymes, have gained intent interest as a target for alternative drug design.

Pondering the current challenges of drug resistance and other dire effects, the second part of my thesis comprehends the investigation of *Staphylococcus aureus* Pdx proteins via the aforementioned biophysical techniques and 3D structural studies, for lining up the Pdx proteins as possible alternative targets. Knowing the differences between the eukaryotic and prokaryotic Pdx1 and Pdx2 sequences with thermodynamic and structural differences in their Pdx complexes assembly. Further, the obtained structural insight into the *P.vivax* Pdx1, and the Pdx complex dynamic oligomerization behavior in solution, *S.aureus* Pdx proteins were selected with a focus on Pdx1 3D structure and the Pdx complex dynamic studies. *Sa*Pdx1, Pdx2, and the Pdx complex were studied by applying complementary bioanalytical techniques, likewise applied in *P.vivax* Pdx proteins investigations. The data obtained by applying dynamic light scattering (DLS), small-angle X-ray scattering (SAXS), and mass photometry (MP) indicate *Sa*Pdx1 in a hexameric state without dodecameric equilibrium form in solution (reported first time in bacteria). *Sa*Pdx2 was found stable and in the monomeric solution state, indicated by SEC, DLS and SAXS data.

Since *S.aureus* Pdx1 was showing a hexameric solution state, therefore it was interesting to see whether the hexameric Pdx1 can form complex after mixing with its Pdx2 counterpart. SEC, DLS, and SAXS data indicate the Pdx complex formation with confirmation of the respective protein subunit bands by gel electrophoresis. Whereas, the low-resolution 3D SAXS data shows a tetrameric Pdx complex composed of Pdx1 and Pdx2 in 1:1.

Interestingly, contrary to the solution state, *Sa*Pdx1 was found in a potential dodecameric form in its 3D structure after investigating its solved crystalline structure. The online interface investigating tool (PISA) which is used for the calculation of the interface score also known as complex formation significance score (CSC), demonstrates a more stable hexamer with less possibility of dodecamer formation for *Sa*Pdx1. However, the presence of PO_4 ions at the hexamer interface of the protomers is one of the possible reasons for a potential dodecameric crystalline state. In contrast to the *P.vivax* and other reported Pdx1 crystal structures, complete helices αN and α2′ are

appeared and solved in *S.aureus* Pdx1 crystal structure. The reported data indicate that these helices appeared after Pdx1 binding to Pdx2 proteins and/or R5P bind to Pdx1 with function in Pdx complex stabilization and enzyme catalysis. The helix αN present at the interface of Pdx1-Pdx2 proteins is a potential target region for designing inhibitors by blocking the associated ammonia tunnel involved in the transport of NH_3 from the Pdx2 subunit active site to the Pdx1 catalytic center.



Graphical Abstract

1.0. Zusammenfassung

Die wichtigste Voraussetzung für die Inhibitor-basierten Selektionsansätze zur Heilung von Krankheiten und Infektionen basiert auf dem Nachweis, dass das Ziel des Erregers nicht im Wirt vorhanden sein darf und/oder der potenzielle Inhibitor in den Stoffwechselwegen des Wirts ungewöhnlich sein muss, um schädliche Auswirkungen auf die Wirtsproteine zu vermeiden. Im Idealfall sollten neue Medikamente entwickelt werden, die speziell gegen den Parasiten wirken und dabei nur minimal oder überhaupt nicht toxisch für den Menschen sind. In diesem Zusammenhang wurde das essentielle Vorhandensein von Vitamin B_6-Biosyntheseenzymen in den Stoffwechselwegen einer Mehrheit pathogener Organismen und ihre Abwesenheit beim Menschen den Pyridoxalphosphat-Enzymkomplex (PLP) als potenzielles Wirkstoffziel erkannt. In diesem Zusammenhang konzentrierten sich meine Forschungsaktivitäten auf die Struktur-Funktions-Analyse des PLP-Synthase-Komplexes. Die Hauptstruktureinheit und das katalytische Zentrum des PLP-Synthase-Komplexes bilden die Untereinheit Pdx1 (PLP-Synthase-Domäne), die einen Dodecamer bildet, und Pdx2 (Glutaminase-Domäne), die unabhängig an Pdx1 bindet und einen 24mer Pdx1-Pdx2 (Pdx) Komplex bildet.

Die wesentliche Bedeutung des PLP-Synthase-Komplexes beruht auf seinen Strukturelementen, die den Eintritt des Substrats zum aktiven Zentrum steuern, sowie auf den Pdx1-Pdx2-Schnittstellenbereichen. Die Nutzung des PLP-Synthase-Enzymkomplexes als potenzielles Wirkstoffziel für Arzneimittelentdeckungsansätze erfordert Informationen über diese wichtigen Strukturelemente und/oder über die Pdx1-Pdx2-Schnittstellenregionen. Das PLP-Synthase-Komplex-Enzym nutzt die gleichen Substrate (Glutamin und Zucker, d.h. Ribose 5-Phosphat und Glyceraldehyd 3-Phosphat) wie einige essentielle Wirtsenzyme z.B. Glutaminamido-Transferase, Ribose-5-Phosphat-Isomerase und Glyceraldehyd 3-Phosphat-Dehydrogenase, die an verschiedenen Stoffwechselwegen beteiligt sind. Daher sind die schnittstellenbasierten Pdx1-Pdx2-Inhibitoren ein besserer Ansatz als das substratbasierte Inhibitordesign, das auf die PLP-Synthase-Enzyme abzielt.

Unter Berücksichtigung des Potenzials von PLP-Synthase-Enzymen für neuartige Arzneimittelentwicklungsstudien konzentriert sich der erste Teil meiner Arbeit auf die Analyse der *Plasmodium vivax* PLP-Syntheseproteine (Pdx1, Pdx2 und der Pdx-

Komplex) mittels biophysikalischer Charakterisierung und strukturellen Untersuchungen. Als Plasmodial-Arten verursachen Protozoen-Erreger schwere und tödliche Malariaerkrankungen und haben einen wirksamen Resistenzmechanismus gegen die meisten der derzeit verfügbaren Malariamedikamente entwickelt. Dies führt zu einer dringenden Notwendigkeit, neue Ziele für Malariamedikamente zu identifizieren. Um die PLP-Enzyme und deren Komplex im Hinblick auf erste Arzneimittelentdeckungsuntersuchungen zu charakterisieren, wurden das *Plasmodium vivax* Pdx1, Pdx2 und der Pdx1-Pdx2 (Pdx) Komplex unter Verwendung komplementärer biophysikalischer und bioanalytischer Techniken wie dynamischer Lichtstreuung (DLS), Massenphotometrie (MP), Kleinwinkel-Röntgenstreuung (SAXS) und Elektronenmikroskopie (EM) untersucht. Darüber hinaus wurden für Pdx1 und den Pdx-Komplex auch Strukturstudien mittels Röntgenkristallographie- durchgeführt, bei der Synchrotronstrahlung am DESY für die Beugung an Ein-Kristallen verwendet wurde.

Die biophysikalischen Untersuchungen ergaben einen dodecamerischen Pdx1 und einen monodispersen Pdx-Komplex. Pdx2 wurde jedoch in monomeren und in verschiedenen oligomeren Zuständen in Lösung identifiziert. Interessanterweise führte das Mischen von oligomerem und polydispersen Pdx2 mit dodecamerischem monodispersen Pdx1, untersucht durch zeitaufgelöste DLS und nanopartikel-Tracking-Analyse (NTA), zum allmählichen Verschwinden des oligomeren Pdx2 und zur Bildung eines monodispersen Pdx-Komplexes innerhalb von Stunden. Die Daten zeigten, wie wichtig die zeitaufgelöste Untersuchung für die Untersuchung der Dynamik eines komplexen Reaktionsprozesses ist. Anschließend ergaben die Elektronenmikroskopie-Studien die Bildung eines ungesättigten Pdx1-Pdx2-Komplexes, dessen Sättigung mit dem oligomeren Pdx2 beobachtet wurde. Darüber hinaus ergaben die Kristallstrukturuntersuchungen auch eine dodecamerische PvPdx1, was darauf hindeutet, dass die dodecamerische Form von Pdx1 explizit ein besonderes Merkmal von Plasmodien ist.

Zweifellos ist die Entdeckung des Breitspektrumantibiotikums Penicillin durch Alexander Fleming eine der größten Errungenschaften in der Geschichte der Medizin, die über Jahrzehnte unzählige Leben auf der ganzen Welt gerettet hat. Aber, die natürliche Resistenzentwicklung in Mikroben aufgrund der übermäßigen Verwendung

der Antibiotika, und langfristigen Exposition gegenüber diesen Antibiotika führen zu einem sehr schnellen Verlust der Wirksamkeit. Aufgrund der sehr schnellen Anpassungsfähigkeit ist der Methicillin-resistente *Staphylococcus aureus* (MRSA), der häufig bei Krankenhausinfektionen auftritt, zur lebensbedrohlichsten Mikrobe geworden. Mit nur wenigen verfügbaren therapeutischen Möglichkeiten hat sich seine Resistenz gegen die derzeit verfügbaren Antibiotika erhöht. Daher kündigten betroffene Institute und auch mehr als 20 führende biopharmazeutische Unternehmen im Juli 2020 den "AMR Action Fund" mit dem Ziel an, bis zum Jahr 2030 zwei bis vier neue Antibiotika oder alternative Therapeutika auf den Markt zu bringen [1]. In Bezug auf die alarmierenden Nebenwirkungen von Antibiotika und Themen wie Multiresistenzen, werden neuartige Ansätze der Ausrichtung auf die Stoffwechselwege der Mikroben, vor allem diejenigen, die beim Menschen fehlen, wie Vitamin B_6 synthetisierende Pdx Komplex-Enzyme, zum Ziel für alternative Arzneimittel-Entwicklungen.

Der zweite Teil meiner Abschlussarbeit, der die aktuellen Herausforderungen der Arzneimittelresistenz und anderer negativer Effekte betrachtet, beinhaltet die Untersuchung von *Staphylococcus aureus* Pdx-Proteinen mit den oben genannten biophysikalischen Techniken und 3D-Strukturstudien, um die Pdx-Proteine als mögliche alternative Ziele zu identifizieren. Dabei wurden die Unterschiede zwischen den eukaryotischen und prokaryotischen Pdx1 und Pdx2-Sequenzen mit thermodynamischen und strukturellen Unterschieden in ihrer Pdx-Komplex-Baugruppe untersucht. Darüber hinaus wurden die erhaltenen strukturellen Einblicke in die *P.vivax* Pdx1 und das dynamische Oligomerisierungsverhalten des Pdx-Komplexes in Lösung untersucht. *S.aureus* Pdx Proteine wurden mit einem Fokus auf der Pdx1 3D-Struktur und der dynamischen Studien der Pdx -Komplexe ausgewählt. *Sa*Pdx1, Pdx2 und der Pdx-Komplex wurden unter Anwendung komplementärer bioanalytischer Techniken untersucht, die ebenfalls in *P.vivax* Pdx-Proteinuntersuchungen angewendet wurden. Die durch die Anwendung dynamischer Lichtstreuung (DLS), der Kleinwinkel-Röntgenstreuung (SAXS) und der Massenphotometrie (MP) gewonnenen Daten deuten auf *Sa*Pdx1 in einem hexamerischen Zustand ohne dodecamerische Gleichgewichtsform in Lösung hin (erstmals bei Bakterien berichtet). *Sa*Pdx2 wurde stabil und im monomeren Lösungszustand gefunden, der durch SEC, DLS, und SAXS-Daten verifiziert wird.

Da *S.aureus* Pdx1 einen hexamerischen Lösungszustand zeigte, war es daher interessant zu sehen, ob das hexamerische Pdx1 nach dem Mischen mit seinem Pdx2-Pendant einen Komplex bilden kann. SEC, DLS, und SAXS-Daten zeigen die Pdx-Komplexbildung mit Bestätigung der jeweiligen Protein-Untereinheitsbanden durch Gelelektrophorese. Während die 3D-SAXS-Daten mit niedriger Auflösung einen tetramerischen Pdx-Komplex zeigen, der aus Pdx1 und Pdx2 im Verhältnis 1:1 besteht.

Interessanterweise wurde *Sa*Pdx1 entgegen dem Zustand in Lösung in der 3D-Struktur in einer potentiellen dodecamerischen Form gefunden, nachdem seine gelöste kristalline Struktur untersucht wurde. Das Online-Schnittstellen-Untersuchungstool (PISA), das für die Berechnung des Schnittstellen-Scores auch als *complex formation significance score* (CSC) bekannt ist, zeigt ein stabileres Hexamer mit geringerer Möglichkeit der Dodecamerbildung für *Sa*Pdx1. Das Vorhandensein von PO_4-Ionen an der Hexamer-Schnittstelle der Protomere ist jedoch einer der möglichen Gründe für einen potentiellen dodecamerischen kristallinen Zustand. Im Gegensatz zu den *P.vivax* und anderen gemeldeten Pdx1-Kristallstrukturen werden die Helices αN and α2´ vollständig in der *S.aureus* Pdx1-Kristallstruktur angezeigt und gelöst. Die beschriebenen Daten deuten darauf hin, dass diese Helices nach Pdx1-Bindung an Pdx2-Proteine und/oder R5P-Bindung an Pdx1 eine Funktion in der Pdx-Komplexstabilisierung und Enzymkatalyse haben. Die an der Schnittstelle von Pdx1-Pdx2-Proteinen vorhandene Helix αN ist eine potenzielle Zielregion für die Entwicklung von Inhibitoren, indem der zugehörige Ammoniaktunnel blockiert wird, der am Transport von NH_3 vom aktiven Zentrum der Pdx2-Untereinheit zum Katalysezentrum Pdx1 beteiligt ist.

2.0. Introduction

2.1. Malaria: Beware of the bite

Malaria infects more than 250 million humans causing up to 0.5 million mortalities annually. The infecting pathogen of malaria is Plasmodium, a protozoan pathogen which is a member of the phylum Apicomplexa. The five plasmodial species affecting humans are *Plasmodium falciparum, Plasmodium vivax, Plasmodium malariae, Plasmodium ovale,* and *Plasmodium knowlesi,* all species together infect millions of individuals and cause hundreds of thousands of deaths around the globe annually [2, 3]. The warfare against malaria as an infectious disease started with the discovery by Ross who discovered the importance of *Anopheles* mosquitoes, and Grassi in 1898 showed that the transmission of the malarial parasites occurred through the infected mosquito bite [4]. World Health Organization (WHO) African Region report an estimated 215 million cases in the year 2019, accounting for 94% of the global cases originate in 87 malaria-endemic countries [5]. In 1955, WHO launched the first time the ‘‘Global Malaria Eradication Programme’’, ended in 1969 with 27 countries declared as malaria-free [6]. The WHO Global Technical Strategy (GTS) 2016-2030 for malaria aimed at a reduction in malaria cases up to 75% by the year 2025, and 90% by the year 2030. In the year 2020, a milestone for several significant health and development goals globally, WHO has decided for much dire need of coordinated efforts to reduce and eliminate malaria wherever and as soon possible [5].

2.1.1. Plasmodium: Invisible Yet Invincible

Plasmodium species being causative agents of malaria have a complex life cycle span comprised of two hosts i.e., the vector host mosquito from the genus *Anopheles* and the second vertebrates host such as humans and rodents, etc. Plasmodium has distinct asexual morphological stages, a span in both hosts named sporozoite and merozoite. Plasmodium sporozoites once injected into humans by the infected mosquito bite, reach out to the liver, where it starts its hepatic life stage by dividing and differentiating into schizonts, comprising thousands of hepatic merozoites. The produced merozoites are afterward released into the blood, thereby initiating their erythrocytic stage of the life cycle through replicating in red blood cells (RBCs). After many stages of multiplication in RBCs, they increase their number and destroy RBCs, thereby developing malarial symptoms of fever and sweating. Subsequently, these asexual parasites then

differentiate into the gametocytes at the final stages in RBCs and ensure parasite transmission after taken up by the mosquito during feeding. In mosquitoes, these cytes mature and develop into gametes and sporozoites in multiple steps [7].

2.1.2. *Plasmodium vivax*: Relapsing plasmodial parasite

Unfortunately, the eradication of malaria is not accomplished so far in most of the underdeveloped countries, despite the coordinated efforts by World Health Organization (WHO) and other collaborative regulatory institutions. Although *P.falciparum* is responsible for the majority of the malarial cases and deaths around the world and especially in Africa. Whereas, *P.vivax* has a wider geographical distribution and is the main causative organism of almost half of malaria cases outside Africa. Further, *P.vivax* exhibits a bit different and complex life cycle in the liver of the mammalian/human host that makes *P.vivax* infections more worrisome. In the liver, some of the asexual sporozoites of *P.vivax* remain in a dormant form called hypnozoites and do not develop into schizonts immediately. These hypnozoites can resume hepatic development at any time at the impulse of still some unknown factors and may result in relapse after weeks, months, or years from the primary exposure, causing repeated clinical episodes of malaria from a single infection [8]. Therefore, the WHO currently prioritizes investigations targeting *P.vivax,* rather than other Plasmodium species [8, 9].

The drug resistance for *P.vivax* and *P.falciparum* has been observed and noted in data collected from all around the world [10] with a consistent growth pattern [3]. This constant developing resistance of the plasmodial parasite against the antimalarial drugs is compelling scientists in shifting the strategies for the development of the novel class of drugs, that may inhibit various key enzymes important in the life cycle of the Plasmodium. The drug targets that interfere with the metabolism of the parasite, like the vital vitamin biosynthesis pathways, are the key focus for the latest antimalarial drug designing and discovery investigations [11].

2.2. Antibiotic emergency: “Bad bugs-no drugs”

The story of antibiotics and their resistance started with the discovery of the very first broad-spectrum antibiotic penicillin in the year 1943 [12]. At that time, Alexander Fleming mentioned in his Nobel Lecture in 1945, that microbes get resistant to penicillin at a certain dose level [13, 14]. Thus, warned and familiarized the community at the same time with new terminology of antimicrobial resistance. In the year 2016, the

United Nations General Assembly (UNGA) meeting, accredited officially the link between antimicrobial resistance (AMR) and antibiotics misuse costing 40 billion dollars per year [15], with an expected increase of 36% in the future. This reveals their importance from both the societal and economic perspectives [16]. According to WHO, infections with antimicrobial-resistant bacteria result in approximately 700,000 human deaths globally per year and are estimated to reach 10 million in 2050 [1] and at a cost of 100 trillion USD [15].

2.2.1. Methicillin-Resistant *Staphylococcus aureus*: A kind of superbug

The extremely high genetic plasticity of prokaryotes allows them to efficiently acclimate to their concerning environmental changes [17]. For all antibiotics classes registered and utilized in the market, at least one resistance mechanism is present in the corresponding bacterial strain and is still further emerging with complications in their mechanism [18].

Since some of the antibiotic-resistant bacterial species are spreading more rapidly and becoming epidemics. Methicillin-Resistant *Staphylococcus aureus* (MRSA), a worrisome pathogen with serious morbidity and mortality [15], is a precedence AMR microbial pathogen. Due to resistance acquiring ability MRSA infections are a major burden on healthcare worldwide, infecting more than 150,000 individuals in the European Union (EU) annually. This causes an extra burden on in-hospital costs of around 380 million euros on EU healthcare systems [19]. Therefore, MRSA infections establish an imperative and still-evolving public health challenge globally, as well as for the European states. Successful controlling strategies of MRSA spread and related infections in some countries offer opportunities for identifying effective interferences and reassessment of best practices to apply [20].

Staphylococcus aureus is a gram-positive bacterial species usually found on the skin, skin glands, and mucous membranes, mouth and nose of humans and some other animals. It can colonize the skin and other soft tissues, bloodstream, and respiratory tract. In systemic infections, it may cause osteomyelitis (bone infection), mastitis (breast tissues inflammation), serious bloodstream septicaemic, and chronic wound infection [20]. Because of its ability to form biofilm on abiotic surfaces and on implant materials, it is therefore responsible for implant-associated infections. *S.aureus* is a toxin and superantigens producer with a wide range of small-sized proteins and other virulence

factors which may initiate a storm of cytokines with the ability to induce toxic shocks and lytic effects on host cells. Further, these toxins associated release cytokines result in toxin-mediated medical conditions called toxic shock syndrome or sepsis-like syndrome [21].

Nevertheless, one of the utmost challenges in the effective treatment of *S.aureus* related infections is its capability of acquiring resistance to several known classes of antibiotics [21]. But still, for long-term success in controlling MRSA infections, novel drug targets with coordinated actions between different healthcare sectors will be of increasing importance [20].

2.3. Pyridoxal 5-phosphate (PLP) synthase: A putative novel drug target

Admittedly, the disastrous emergence of drug resistance in prokaryotic as well as eukaryotic human pathogens, demands an urgent need and continuous search to discover and develop novel chemotherapeutics [22]. Ideally, the novel drug development strategy should typically be based on targeting the pathogen/s specifically, with least or almost no possible toxicity to the human host [23]. Therefore, based on drug design, the drug targets would be distinctly different from the host's metabolic processes or even their absence in the host [24]. PLP is synthesized via the *de-novo* vitamin B_6 biosynthesis pathway reported in bacteria, Plasmodium, fungi, and plants. Its absence in mammals makes it an admirable drug target to be adequately investigated in terms of structural studies utilizing the most advanced and sophisticated X-rays synchrotron source at DESY.

2.3.1. Pyridoxal 5-phosphate: vitamin B_6

In living cells, for a successful enzyme reaction in many biological processes, the enzyme/s are assisted by some nonprotein components, called coenzymes or cofactors. These cofactors are not translated or synthesized by the ribosome, rather are formed by some other enzymes, being utilized in other enzymatic reactions. My research is focused in pertinent detail on one of such enzymes i.e. PLP synthase complex that synthesizes the co-factor Vitamin B_6.

Vitamin B_6 was first time reported in 1932 [25]. Its descriptive name and structure as Pyridoxine were published first in 1939 and found similar to pyridine ring structure [26]. Its synthesis was also reported in the same year [27]. The current Recommended Dietary Allowance (RDS) for vitamin B_6 set by the American National Institute of

Health (NIH) is 2.0 milligrams per day with an upward tolerance of 100 mg per day for adults [28, 29].

Vitamin B_6 in active pyridoxal 5-phosphate (PLP) form carry out various functions in many enzymatic reactions with a majority of them catalyze amino acid metabolism reactions and function as a transient carrier of specific functionalities in a variety of metabolic processes. PLP-dependent reactions like amino acids decarboxylations, α, β-elimination, racemizations, and transaminations [30, 31], gluconeogenesis and lipid metabolism [32-34] have been extensively reviewed. PLP catalyzed reactions additionally include elimination or replacement of an electrophilic group, and phosphorolysis [33, 35]. The active PLP, as an essential cofactor, also participates in the catalysis of various enzymes including isomerases, hydrolases, transferases, oxidoreductases, and lyases [30]. It also participates in reactions like haemoglobin biosynthesis, the formation of several amines actively involved in neurotransmissions, and neuronal-based reactions e.g. involved in epinephrine, serotonin or norepinephrine, and sphingolipids synthesis [30, 31]. Besides its role as a cofactor, it exhibits an antioxidative nature against reactive oxygen free radicals and participates in many radical-mediated reactions [36], in amino-sugar deoxygenation [37], and glycogen breakdown in muscles utilizing glycogen phosphorylase B [35].

Regarding the structure and other isomeric forms, vitamin B_6 is composed of six vitamers. All vitamers have a common pyridinium ring structure, divided into two groups (a) non-phosphorylated forms which include pyridoxamine (PM), pyridoxine (PN), and pyridoxal (PL), and (b) their respective phosphorylated forms which comprise pyridoxamine 5-phosphate (PMP), pyridoxine 5-phosphate (PNP) and pyridoxal 5-phosphate (PLP). The phosphorylated vitamer forms PLP and PMP represent the biologically active forms and the remaining forms are the PLP precursors [30]. The chemical structures of these biologically interconvertible vitamers are shown in **figure 1**.

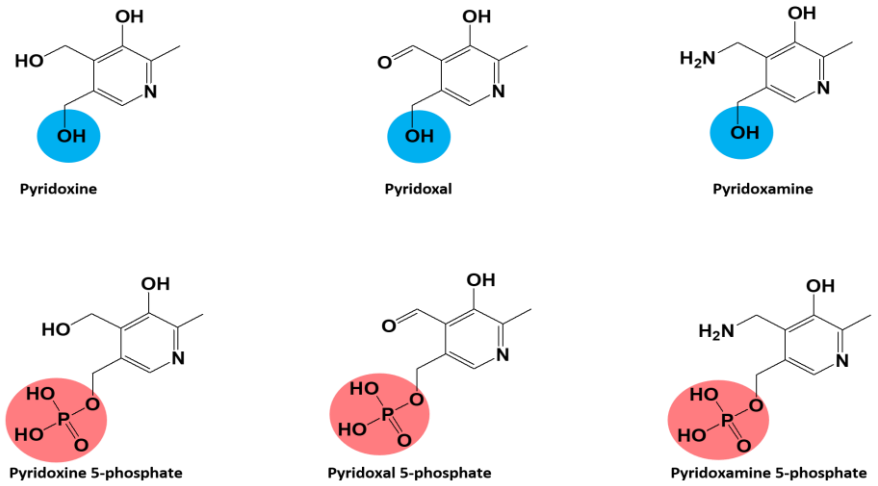


Figure-1. Vitamin-B_6 non-phosphorylated and phosphorylated vitamers.

Interconvertible non-phosphorylated pyridoxine, pyridoxamine, pyridoxal vitamer, and their phosphorylated forms i.e. pyridoxine 5-phosphate (PNP), pyridoxamine 5-phosphate (PMP), and pyridoxal 5-phosphate (PLP).

2.4. Pyridoxal 5-phosphate biosynthesis

2.4.1. *De novo* PLP Biosynthetic Pathways

Enzymes for the *de novo* vitamin B_6 biosynthesis are reported in some organisms, i.e., archaebacteria, bacteria, protists, fungi, and plants [38, 39], whereas humans obtain it from their diet. Vitamin B_6 in the phosphorylated form (PLP or PNP) can be produced directly by two distinct *de novo* biosynthesis pathways i.e. deoxy-xylose 5-phosphate (DXP)-dependent and DXP independent (ribose 5-phosphate or R5P) pathways. Both routes comprise highly conserved genes, but the relative comparison also shows that they are diverse in their nucleotides as well as their translated amino acid sequences. This suggests that both the DXP dependent and independent pathways diverged during the course of evolution.

2.4.1.1. DXP-dependent pathway

The DXP-dependent pathway of vitamin B_6 biosynthesis represents the first pathway uniquely identified and studied extensively. It was first reported in a subset of the γ-division of proteobacteria and was further studied in the gram-negative bacterium, *Escherichia coli* in detail [40]. Despite the original speculation, this pathway is not widely distributed. Additionally, it is astonishingly a very complex pathway and requiring six different enzymes i.e. GapA, PdxB, PdxF, PdxA, DXPS, PdxJ involved in a complicated mechanism for PNP production. Additionally, it depends upon PdxH for the synthesis of the active PLP [38, 41-44].

In a detailed mechanism, the DXP pathway is divided into two branches/steps. One is involved in the synthesis of 4-phosphohydroxy L-threonine (4-HTP). The second step is required for the deoxyxylulose 5-phosphate (DXP) synthesis and both converge in the last step to form the phosphorylated PNP [38]. The starting substrates for this pathway are D-erythrose 4-phosphate (pentose phosphate pathway) utilized in one step, and glyceraldehyde 3-phosphate (glycolysis) and pyruvate (glycolysis) are needed in the second step to synthesize DXP [38].

The initial step requires the oxidation of D-erythrose 4-phosphate by enzyme erythrose 4-phosphate dehydrogenase **(GapA)** and produce D-erythronate 4-phosphate. The next step covers the oxidation of the D-erythronate 4-phosphate by the enzyme D-erythronate 4-phosphate dehydrogenase **(PdxB)** requiring cofactor NAD to produce (3R)-3-hydroxy-2-oxo-4-phosphonooxybutanoate [45]. In the next step, an enzyme **PdxF** (phosphoserine aminotransferase) catalyzes the PLP-dependent transamination between (3R)-3-hydroxy-2-oxo-4-phosphono-oxybutanoate and glutamate, forming 4-hydroxy-L-threonine phosphate **(HTP)** and eliminating 2-oxoglutarate [46, 47]. **PdxA** (de-hydrogenase) catalyzes **HTP** decarboxylation**,** eliminating CO_2, and producing 3-hydroxy-1-aminoacetone phosphate **(AAP)** (the end product of this step and is utilized in the condensation with the DXP (from the second step) [41, 43, 48].

In the second route, DXP is synthesized from the simplest alpha-keto acid (pyruvate) and G3P by the enzyme Deoxy-D-xylulose 5-phosphate synthase **(DXPS)**. DXPS enzyme has been characterized in detail both biochemically and structurally [49]. The process of PNP synthesis is finalized by the enzyme PNP synthase **(PdxJ)**, which catalyzes the cyclization of AAP and DXP in condensation reaction [50-52]. PNP can also be oxidized to PLP by the enzyme PNP oxidase **(PdxH)** in the salvage pathway [53]. PdxH requires co-factor flavin mononucleotide (FMN) and can use both PNP and PMP as substrates in the synthesis of PLP [54, 55], details are presented in **figure 2**.

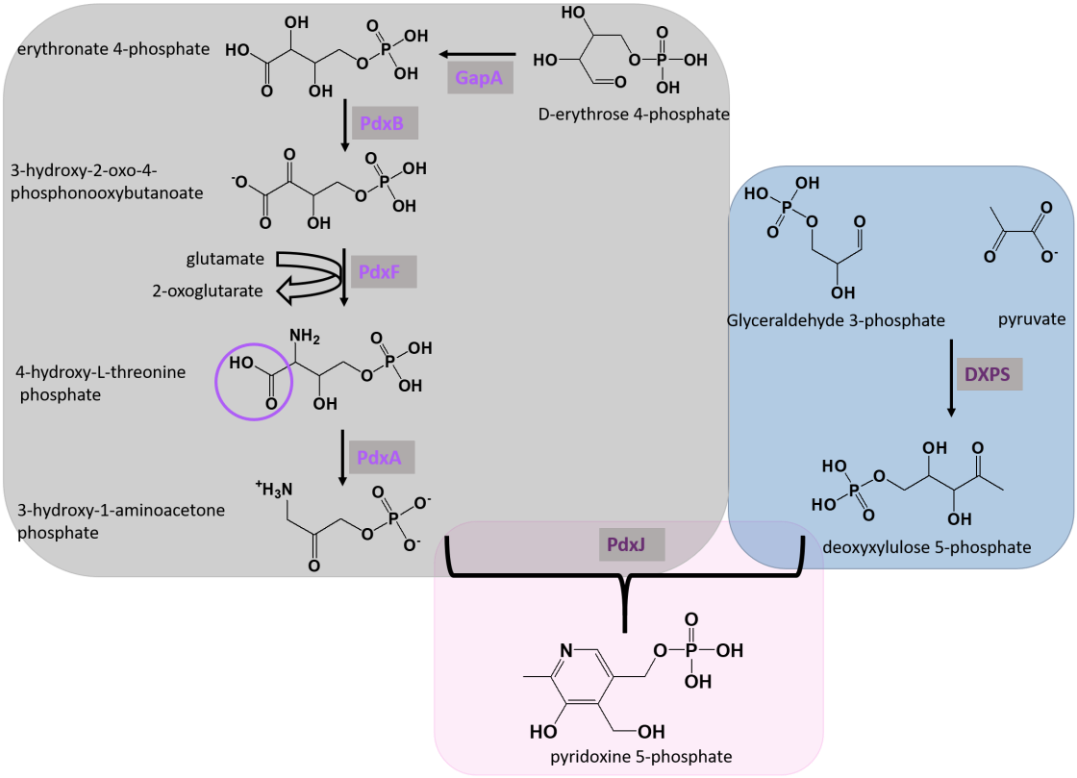


Figure-2. DXP dependent metabolic pathways of vitamin B6 biosynthesis.

DXP dependent pathway consists of two branches having final products 3-hydroxy-1-aminoacetone phosphate in one route and deoxyxylulose 5-phosphate in the second route, condensing to pyridoxine 5-phosphate (PNP) by PNP synthase enzyme (PdxJ).

2.4.1.2. DXP-independent (R5P-dependent) pathway

The DXP-independent pathway was first accidentally identified in *Cercospora nicotianae* while exploring the resistance mechanism of the fungus against the oxidative stress [56, 57]. Later, it was found that this pathway is present and dispersed across all four kingdoms of life i.e. bacteria, archaea as well as in eukaryotes including fungi and plants [56, 58-63]. The unforeseen discovery of this pathway qualifies a few lines to write. The SOR1 (singlet oxygen resistance 1) gene found in fungus (*Cercospora nicotianae*) was revealed to be involved in the singlet oxygen resistance [56, 64]. The mutation that happens in this gene revealed its importance in vitamin B_6 biosynthesis essentially [56]. Though this gene is present and conserved in four kingdoms of life [65] but does not show any homology to the formerly identified vitamin B_6 biosynthetic genes. Thus, suggesting that the SOR1 gene is involved in a novel vitamin B_6 biosynthesis pathway.

Later, it was revealed that the gene (*pdx1*) served as an oxygen radical scavenger, is also required for PLP synthesis [56, 65]. The coded protein SOR1 (which was later named

Pdx1) in a complex with a second protein (called Pdx2), which provides the heterocyclic nitrogen to the PLP ring, are involved in the PLP synthesis [58]. This led to the discovery of a second gene *pdx2* linked to *sor1* involved in the "novel" PLP biosynthesis pathway [57]. Sequence alignments classified the protein Pdx2 (encoded by gene *pdx2*) to the family of Triad GATase, which function as transaminase [65]. Later, it was also determined in yeast that ribose 5-phosphate (R5P) and glyceraldehyde 3-phosphate (G3P) are the synthase acceptor substrates [66, 67], and glutamine is the transaminase substrate of the GAT family enzyme [59]. Further, it was established and reported in a series of investigations that Pdx1 in complex with Pdx2 catalyzes the condensation of glutamine, R5P and G3P to form the product PLP [61, 62, 67, 68].

In the detailed mechanism of GATase, Pdx2 catalyzes the hydrolysis of glutamine utilizing the catalytic triad residues Glu-His-Cys to produce neutral ammonia (NH_3). This ammonia moves through a hydrophobic ammonia channel [69] and reaches the active site of Pdx1, where PLP is synthesized in series of reactions from R5P and G3P [70]. Both Pdx1 and Pdx2 proteins constitute a transient complex that is thermodynamically stabilized by the presence of the Pdx2 substrate glutamine. For Pdx complex formation, the ratio of Pdx1 and Pdx2 subunits for optimal catalytic activity was found 1:1 [68]. Upon addition of the substrates R5P and glutamine, a chromophoric intermediate was observed and next the addition of G3P converted the intermediate into PLP [70], details are presented in **figure 3**. In terms of mechanism and reaction intermediates of PLP biosynthesis by ribose 5-phosphate dependent pathway, much remains to be done to illustrate the details.

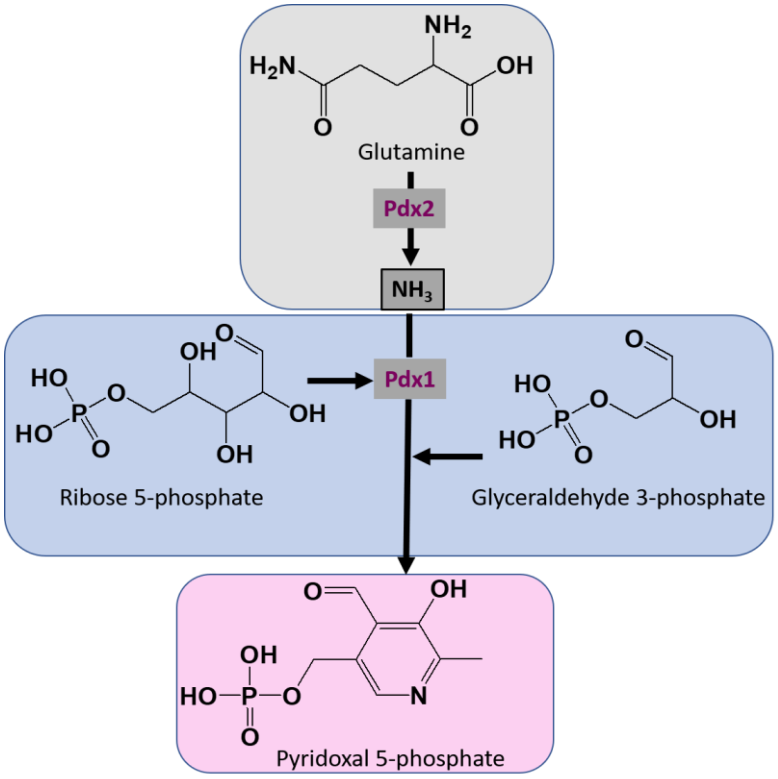


Figure-3. Ribose 5-phosphate dependent pathway.
PLP synthase complex contains two subunits, i.e. glutaminase (Pdx2) and synthase (Pdx1). The glutaminase domain consists of a conserved triad active site required for deamination of the glutamine. Glutamate, NH_3, and water are the products of glutamine hydrolysis. The produced neutral NH_3 is transferred to the active site of the synthase (Pdx1) subunit through a hydrophobic tunnel, where it reacts with R5P (synthase substrate) and forms the intermediate (I_{320}). The intermediate is afterward converted into PLP in a series of reactions at the synthase domain after utilizing G3P.

2.4.2. The Salvage Pathway

Accompanying the direct synthesis of the *de novo* PLP, it can be produced from its precursors in the form of PN, PM, PL as well as from phosphorylated forms of PMP and PNP in the salvage pathway. The details of this pathway are presented in **figure 4**. This pathway functions through the recycling of the B_6 vitamers governed by different Kinases and Phosphatases [71-73], Transaminases, and Oxidases present in the proteome of the reported organisms [74]. The salvage pathway has been studied in detail in *E.coli*, showing two different Kinases that can phosphorylate the vitamers PM, PN, and PL to their respective 5-phosphate forms. Both the Kinases diverge in their substrate specificities i.e. Kinase PdxY acts on PL, whereas Kinase PdxK utilizes the three non-phosphorylated forms as substrates and converts them to their respective phosphate classes. Whereas in eukaryotes, a single Kinase has been reported, and the

crystalline structure of the dimeric Kinase is available in the literature, reported from various organisms [75-77]. Briefly, each of the two protomers contains an active site utilizing nucleotide ATP and metal ions for the enzyme activity [78]. Contrary to the Kinases, a single Oxidase, PdxH, is observed in *E. coli* oxidizes the phosphorylated forms (PMP and PNP) to the active PLP form [54, 79]. The crystal structure of the pyridoxine oxidase functions as a dimer is reported [80, 81].

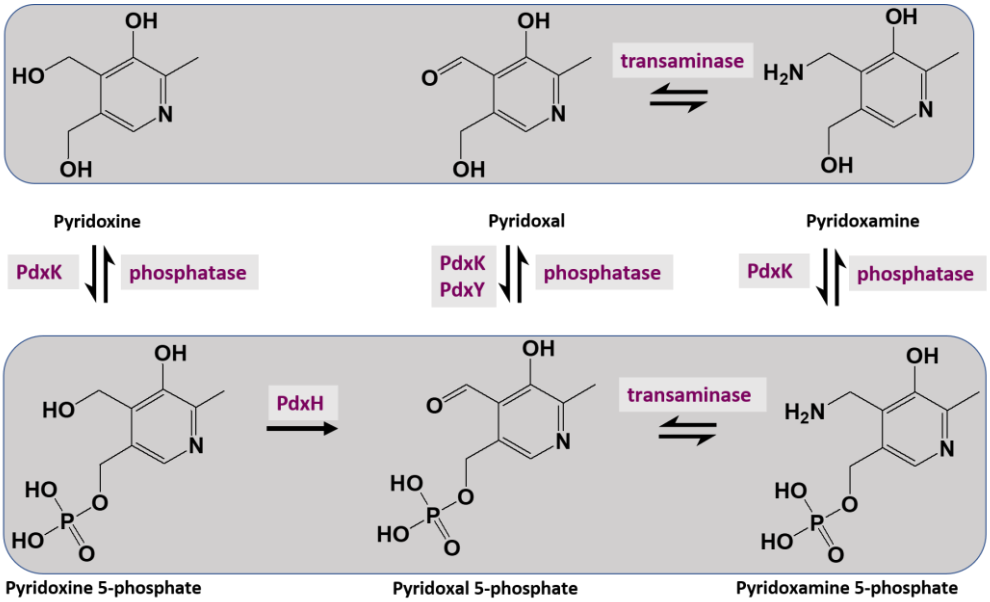


Figure-4. The Salvage pathway of vitamin B_6 biosynthesis.

Vitamin B_6 vitamer interconversion by a salvage pathway is shown here utilizing different enzymes like Kinases (PdxK, PdxY), Phosphatases, Oxidases (PdxH), and Transaminases.

2.5. The Glutamine Amido-Transferases (GATases)

GATases are the biosynthetic enzymes able to catalyze the removal of the amino group in the form of neutral ammonia from glutamine and transform a variety of substrates to novel classes of biomolecules, thereby linking carbon and nitrogen metabolism [82]. Besides, the deaminated product glutamate can be converted to α-ketoglutarate which is a key intermediate in the Krebs cycle, therefore maintaining and balance the cell metabolism [83]. Most of the biological processes involved in the biosynthesis of amino acids, coenzymes, and nucleotides are catalyzed by these glutamine amidotransferases (GATases) utilizing glutamine as a source of ammonia. They consist of two main structural domains; a glutaminase domain, which produces ammonia from the deamination of glutamine, and a synthase domain, which accepts and utilizes the nascent produced ammonia and ultimately produces various other biologically significant biomolecules or co-factors involved in different metabolic processes.

There are two types of GATases i.e. class I or trpG-type or triad GATases and class II or purF-type or N-terminal nucleophile GATases [84]. Both classes of enzymes possess a similar mechanism in terms of their catalysis utilizing a conserved cysteine residue present in the glutaminase binding site and differ in the way of using the nucleophilic cysteine. GATases of class I utilize conserved Cys-His-Glu as an active site, where histidine and glutamate amino acids activate cysteine residues with two conserved glycines and glutamate involved in oxyanion hole formation [69]. In class II GATases the catalytic cysteine residue is located at the N-terminus and two other conserved residues Asn and Gly form an oxyanion hole for stabilization of the intermediate [84].

The PLP biosynthesis proteins i.e. Pdx1 and Pdx2 assemble into class I glutamine amidotransferase enzyme with the former as a PLP synthase domain and the latter function as the glutaminase domain [68, 85]. The ammonia produced in the transamination reaction at the glutaminase domain is afterward transferred through the hydrophobic NH_3 tunnel to the synthase (Pdx1) part, where it reacts with the other synthase substrates resulting in the formation of the product PLP [86]. Distance between the distinct active sites of the synthase and glutaminase domains ranges from 10 Å to 40 Å, a prerequisite for the safe channeling of the neutral ammonia [84]. Allosteric regulation and proper synchronization of the synthase (Pdx1) and glutaminase (Pdx2) active sites play an important role in the mechanism of action of GATase [87], ensuring all the three substrates are consumed in proper order and at the same rate. The reported putative pattern for this synchronization is the Pdx2 dependence on the presence of Pdx1 and Pdx2 substrates and the formation of the intact proteins in the complex [68, 88]. The invariant catalytic triad of cysteine, histidine, and glutamate is located at one end of the core β-sheet of Pdx2 and the interface of both proteins [89].

2.5.1. NH_3 channeling in the β/α-barrel of GATases

The ammonia produced from the deamination of glutamine by Pdx2 is directed via a putative ammonia tunnel that runs through the core of the Pdx1 $(\beta/\alpha)_8$-barrel and reaches the active site of Pdx1, which is presumed to be inaccessible to the solvent molecules. Further, the ammonia in this journey from glutaminase to synthase active site must be protected from conversion to amide or ammonium ions [87]. The opening and closing of the hydrophobic tunnel and the path of ammonia transfer are still

speculated and left for propositions, because of the structural analysis limitation in the Pdx1 domain.

However, in structural investigations, it was proposed that methionine residues control the passage of ammonia [90]. In the kinetic assays with Pdx1 alone, it was confirmed that Pdx1 was functional in the presence of ammonium salts, producing the product PLP in the absence of Pdx2. In addition to this, the kinetic studies have also indicated that the deletion of N-terminal helix (αN) of synthase (Pdx1), indispensable for activation of glutaminase domain, does not affect the PLP synthesis in ammonia-dependent reactions and in absence of Pdx2. The helix αN, which caps the bottom of the Pdx1 hydrophobic β-barrel, functions as an ammonia gate and is located at the interface of both proteins. It encloses the appropriate hydrophobic environment for the safe passage of the reactive neutral ammonia [90, 91]. Later, in plasmodial Pdx proteins, based on structural studies, the role of methionine residues located in the helix αN and β-barrel was proposed in the hydrophobic ammonia tunnel. In Plasmodium, both Pdx1 and Pdx2 active sites are 20 Å apart [69].

2.6. PLPS (PLP synthase or Pdx complex)

Pdx complex consists of two subunits i.e. Pdx1 (synthase subunit) and Pdx2 (glutaminase subunit) and homologs of both Pdx1 and Pdx2 have been identified and reported in various organisms [56, 57, 69, 90]. This thesis probes Pdx1, Pdx2, and the PLP synthase complex from *P.vivax* and *S.aureus*. So, the Pdx1 and Pdx2 nomenclature and their amino acid sequence numbering will be used throughout.

Pdx complex is reported from both prokaryote and eukaryote homologs [69, 92]. It is a D_6-symmetric complex with 12 Pdx1 subunits (Pdx1 = 33 kDa each) and 12 Pdx2 subunits (Pdx2 = 25 kDa each) (Figure 5). The 3D structure of the Pdx complex from the bacteria *Thermotoga maritima* with PDB code: 2ISS [92] and *Bacillus subtilis* (PDB code: 2NV2) [90] and in eukaryotes the chimeric Pdx complex from *Plasmodium berghei* (Pdx1 subunit) and *Plasmodium falciparum* (Pdx2 subunit) [69] has been studied. The reported Pdx complex is a 24-meric giant multienzyme complex, consisting of a dodecameric Pdx1 core that interacts independently with one Pdx2 subunit [90]. The interior of the Pdx1 dodecamer is saturated with the active sites of 12 Pdx1 protomers and 12 Pdx2 subunits studded at the exterior of the Pdx1 dodecamer with their respective active site, facing the peripheral end of a Pdx1 dodecameric barrel

[93]. The active sites of both Pdx1 and Pdx2 are distant from each other with approximately 10 Å – 40 Å. Further, the Pdx2 hydrolyse glutamine to glutamate and producing NH_3, which is later transferred through this distance of 10 Å – 40 Å. It starts its travel from the Pdx2 active site and continues through the hydrophobic interior of the Pdx1 (β/α)8-barrel lined by methionine residues, and reaches the Pdx1 active site [90].

Thermodynamic and structural characterizations revealed that prokaryotic and eukaryotic Pdx proteins and their complexes have substantial structural differences [94]. Additionally, the thermodynamic and biophysical characterization of the Pdx1-Pdx2 complex from *P.falciparum* and *B.subtilis* revealed differences in both complex structures [95]. The Pdx complex is more stable in Plasmodium than in Bacillus [94] with an additional aggregation tendency of *P.falciparum* Pdx complex in solution [69]. Whereas, Pdx complex from *P.vivax,* investigated through different biophysical techniques and X-ray studies, reveals a solution stable Pdx complex [96].

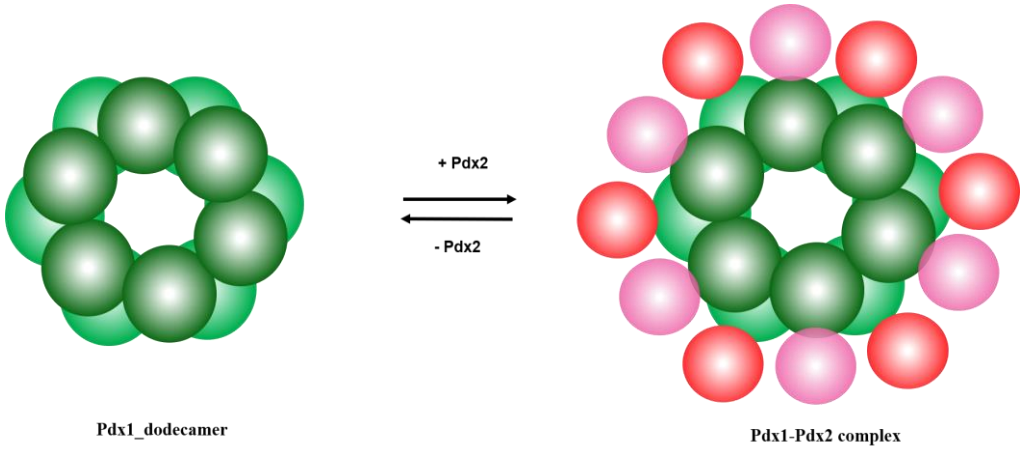


Figure-5. Schematic representation of Pdx1 dodecamer and Pdx complex 24-mer.
Catalytic Pdx1 subunit (light and dark green) forms a cylindrical dodecamer from two hexamers. Each Pdx1 molecule binds to four neighbors Pdx1 molecules. The Pdx2 subunit (light and light red) binds to the Pdx1 subunit like hoops, autonomous of another Pdx2 and establishing the 24meric PLP synthase complex.

2.6.1. Pdx1 (Synthase domain)

2.6.1.1. Pdx1 oligomeric state

Pdx1 forms a cylindrical dodecamer consisting of two opposing interdigitating hexameric rings that stack to each other and dimerize, thus establishing the overall shape of the dodecameric sphere **(Figure 6a,b)** [90, 97]. To date, Pdx1 X-ray crystallographic structures have been reported for many prokaryotes and eukaryotes, mainly in dodecamer states with few exceptions of a hexamer. Pdx1 dodecamer has

been reported from *Bacillus subtilis* (pdb code: 2NV1) [90], *Geobacillus stearothermophilus* (1ZNN) [97], (4WXZ) [93], *Mycobacterium tuberculosis* (4JDY) [98], *Plasmodium berghei* (4ADT) [69] and *Arabidopsis thaliana* (5LNS) [39]. Whereas Pdx1 from yeast *Saccharomyces cerevisiae* (3O06) [91] and (3FEM) [99], and the archaeon *Pyrococcus horikoshii* (4FIQ) [100] crystallizes as hexamer state.

The dodecameric nature of Pdx1 in solution has been characterized through various approaches like dynamic light scattering, SEC, and analytical ultracentrifugation, electron microscopy, and solution scattering (SAXS) investigations [69, 96]. Analytical size investigations demonstrate the *Pyrococcus horikoshii* PdxS hexamer in solution. In crystal structure studies, the asymmetric unit of the PdxS contains a hexamer. The close-up structure investigations show that a long insertion of 37 amino acid residues found in the PdxS sequence from the *P.horikoshii*, seems to be the reason for its hexamer form. Further, the *P.horikoshii* PdxS structure superposition to the other reported dodecameric Pdx1 proteins demonstrates that the additional insertion of 37 amino acids appeared to discomfit the dodecamer formation [100].

The eukaryote *S.cerevisiae*, Pdx1 (Snz1) exists in hexamer form, both in the solution as well as in the crystalline state [99]. Further structural investigations showed that the insertion of Lys177 residue in the Snz1 located at the hexamer interface is the reason for a dodecamer to hexamer shift [99]. In additional independent analytical ultra-centrifugation experiments with PdxS from *Geobacillus stearothermophilus* and Pdx1 *Bacillus subtilis* revealed their existence in a hexamer-dodecamer equilibrium in solution [90, 97]. The hexamer-dodecamer equilibrium in *G.stearothermophilus* PdxS was found dependent on the phosphate and sulfate salts in a salt concentration-dependent manner [97]. Following this, it was also observed in the crystal structure that the sulfate ions are involved in the salt bridges with the conserved residues present on the hexamer-hexamer interface of *G.stearothermophilus* PdxS, aiding in the dodecamer formation [97].

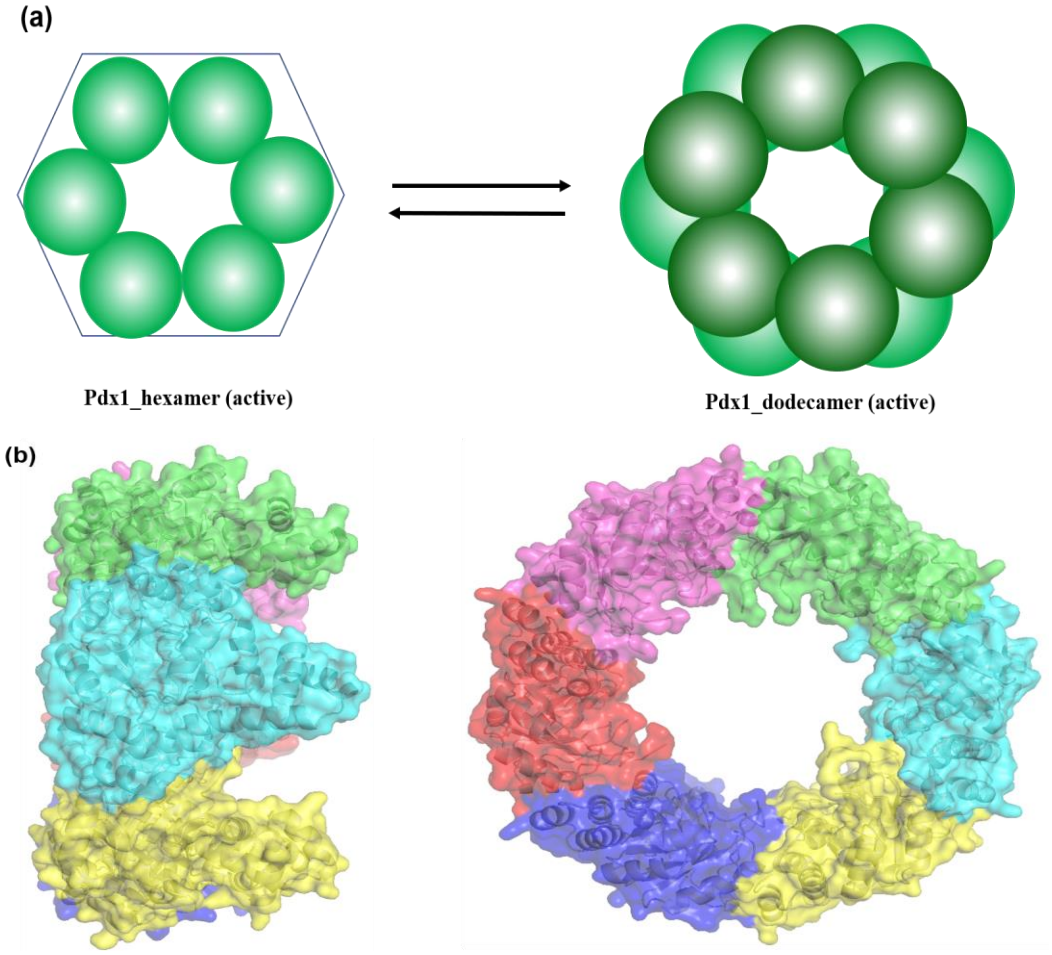


Figure-6. Schematic representation and organization of Pdx1 multimers.
(a) Pdx1 hexamer and dodecamer establish equilibrium in solution with Pdx1 subunits, light blue in one hexamer, forming a dodecamer with a second hexamer in dark blue. **(b)** Pdx1 hexamer side view on the left and dodecamer on the right with front view indicating a hollow and cylindrical sphere, each chain is shown in a different color.

2.6.1.2. Pdx1 (β/α)8 or TIM barrel architecture

Pdx1 is composed of two active sites referred to as P1 and P2 sites, both coordinate the series of enzymatic reactions resulting in the product PLP form. The reaction begins with the R5P binding at the P1 site of Pdx1, covalently to the active site residue lysine. Then the incorporation of ammonia from Pdx2 results in intermediate I_{320} formation. The reaction proceeds further after G3P addition, taking place at the P2 site, where PLP (end product) is observed bound to a second lysine [101].

In zoom-in view, Pdx1 protomer possesses a classic $(\beta/\alpha)_8$ or TIM (Triosephosphate Isomerase) barrel fold consisting of an inner core of 8 parallel β-strands that alternate with 8 α-helices around the β-strands (shown in Figure 7) [97]. The Pdx1 active site in

each protomer is located at the C-terminus of the β-barrel common in $(\beta/\alpha)_8$-barrel enzymes [39, 87, 93]. Moreover, Pdx1 helices are adopted to some prominent modifications participating in the interactions with adjacent protomers needed for multimers especially dodecamer formation e.g. α6′, and α6″. The helical insertion α2′ is aiding to Pdx1 enzymatic activity by controlling and regulating the substrate entry to the active site. Further, the inserted helices α8′ and α8″, which are non-canonical to the classic $(\beta/\alpha)_8$-barrel, decorating the C-terminus of Pdx1 which is involved in the catalysis. The details of these structurally and functionally important helices with insertions are explained in the following paragraphs.

The dodecameric assembly of Pdx1 is facilitated by helix α6 and its two additional helical elongations i.e. helices α6′ and α6″, protruding from the $(\beta/\alpha)_8$-barrel architecture [90, 92, 97]. This helix and its insertions are located at the hexamer interface and are involved in hexamer dimerization through many inter-hexamer interchain contacts [97]. One example indicating the importance of these helices can be observed in the *S.cerevisiae* Pdx1 (Snz1) structure, which exists in hexamer form in solution, as well as in crystalline form. Here the insertion of one extra residue i.e. Lys177 in the sequence between helices α6 and α6′ prevents the dodecamer formation [99]. This small (one residue) insertion induces a spatial conformational alteration in the elongated α6 helix and its insertion helices (α6′, α6″) region when compared to the bacterial homolog *B.subtilis* Pdx1, tempting a steric clash at the hexamers interface [99]. In another situation, where Pdx1 from *P.horikoshii* is reported in hexamer form in the crystalline state because of an inserted region of 37 amino acids in α6′, α6″, preventing the close spatial contact of these helices [100]. Additionally, the hexamer-hexamer interface also consists of hydrogen-bond interactions between the helices α4, α5, α6, α6′, and α6″ regions, aiding to dodecamer formation [90, 92, 97]. Whereas the additional insertion in *P.horikoshii* Pdx1 contributes to the loss of the complementarity on the hexamers interface, a pre-requisite of dodecamer formation [100].

Inter-subunit contacts between the adjacent $(\beta/\alpha)_8$-barrels within one hexamer are mainly established by a bulge created by helix α8″ at the C terminus running parallel to helix α8. Further, helices α2, α3, and α6, loop regions α2′-α2, β4-α4, β3, and the flexible C-terminal regions are equally involved in interactions between the adjacent subunits in a hexamer. The C-terminus is positioned close to the active site of the $(\beta/\alpha)_8$-barrel that

faces the center of the ring-like structure [39], [87]. Aiding to this, the C-terminal tail of 14-28 amino acids is reported to be disordered in all the published native Pdx1 structures [87, 97].

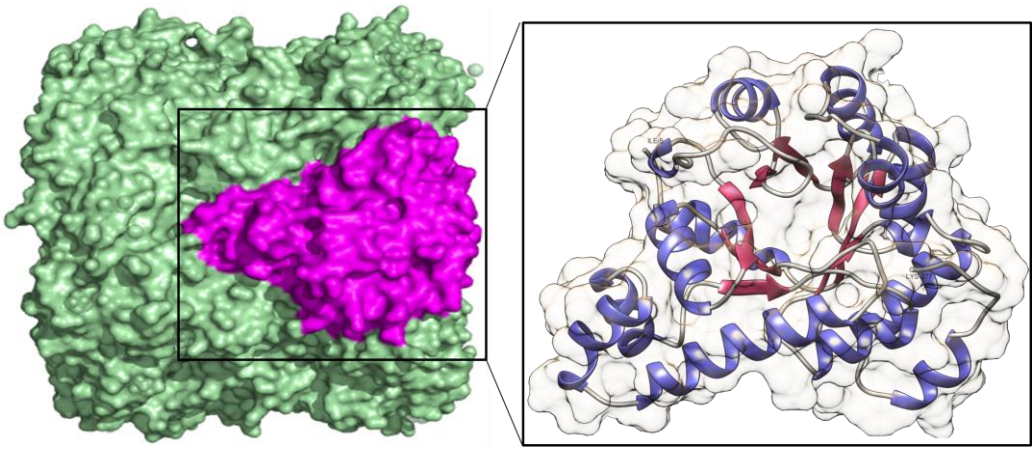


Figure-7. Pdx1 dodecamer surface model with a monomeric TIM-barrel.
Side view of Pdx1 dodecamer in green surface representation with a monomer shown in magenta. To the right, monomeric Pdx1 $(\beta/\alpha)_8$-barrel with secondary structure elements indicated in different colors (helices are shown in blue, strands in purple and loops in light green color), the figure is prepared by chimera [102].

2.6.2. Pdx2 (PLP glutaminase domain)

Pdx2 binds to the N-terminal side of the $(\beta/\alpha)_8$-barrel fold of Pdx1, involved in the formation of Pdx complex (glutamine amidotransferase) [97]. Pdx2 is the glutaminase domain of class-I glutamine amidotransferase, also called the Triad glutamine amidotransferase (GATase) [85, 103]. The class-I GATases have a characteristic Glu-His-Cys catalytic triad. The conserved catalytic triad residues of Pdx2 are reported from *Plasmodium falciparum* with their positions in the sequence i.e. Cys87, His196, and Glu198 [103]. The sequence homology and topology information are shown in **figure 8a-c**, which enlighten that Pdx2 triad domain residues are much conserved throughout the evolutionary process. These triad residues (CHE) in *P.vivax* are located in the helix α5, helix α7 and the end region of loop β13-α7 (**Figure 8 a,b**).

Topology and 3D structural investigations unveil that the main core of Pdx2 is comprised of central β-strands that are predominantly parallel, which create a central β-sheet structure [85, 103]. The structural elements of *P.vivax* Pdx2 in 3D are shown in **figure 9**. The structure illustrates that Pdx2 is an αβα three-layered sandwich structure with Rossmann topology where beta strands form an extended β-sheet through hydrogen bonding, and α-helices surround both faces of the sheet producing a three-

layered sandwiched structure. Overall Pdx2 structure comprises thirteen β-strands and in total nine helices including one α6′ insertion **(Figure 8a,c** and **Figure 9)** following the characteristics of class-I amidotransferases [90].

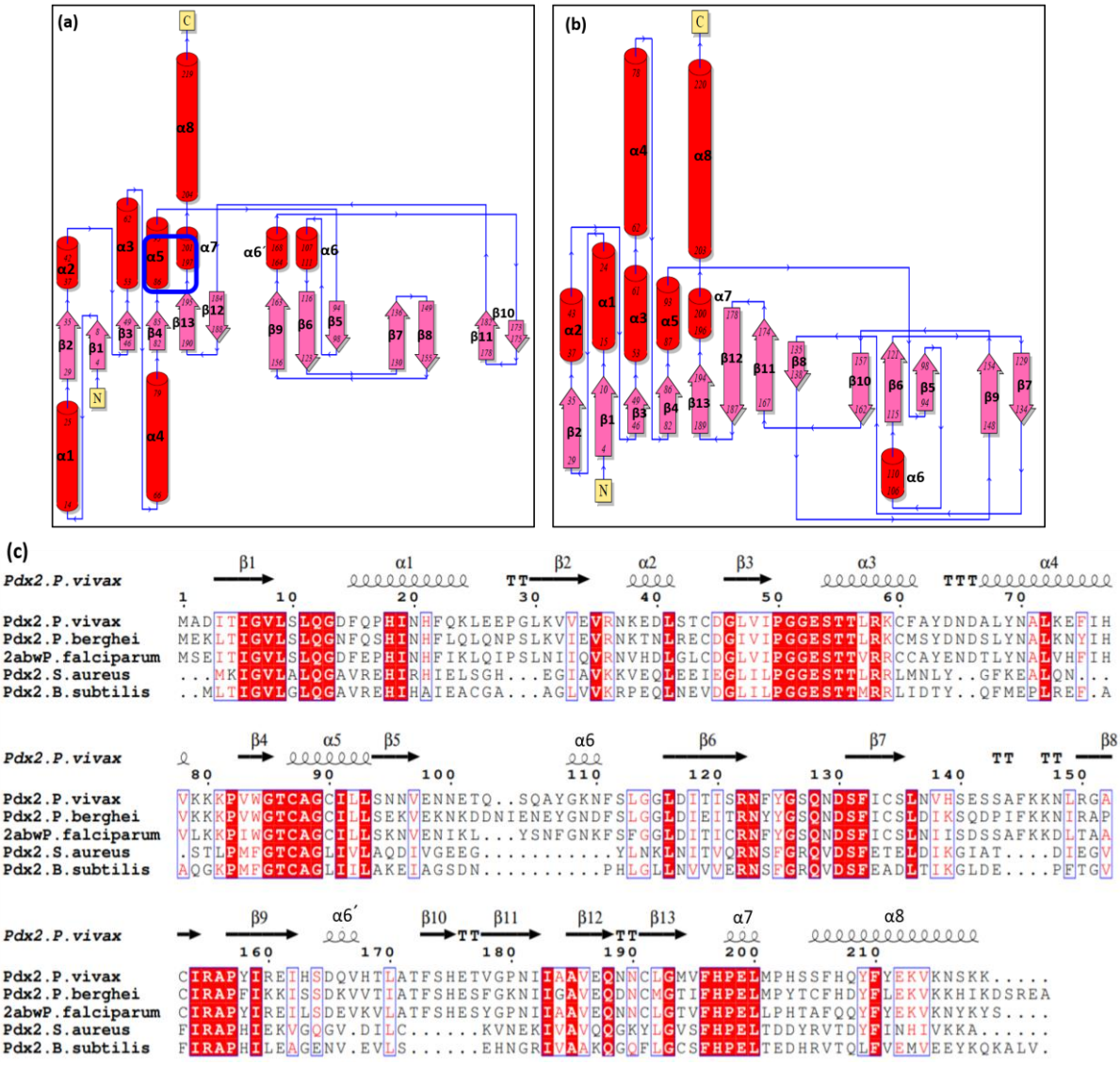


Figure-8. Topology diagram and sequence alignment of *P.vivax* Pdx2 and homologs.
(a) *P.vivax* Pdx2 **(b)** *P.falciparum* Pdx2 (PDB id: 2ABW) structural elements comparison in 2D topology representation. The N and C termini with secondary structural elements are labeled, and residue numbers present at the beginning and end of each structural element are mentioned. Figure **a** and **b** are generated via the online server PdbSum [104]. The catalytic triad Glu-His-Cys domain in *P.vivax* is encircled (blue) in (a). **(c)** Sequence alignment of Pdx2 amino acid sequences from *P. vivax* (A0A1G4GXI4) *P.berghei* (Q4PJX5), *P.falciparum* (Q8IIK4), *S.aureus* (Q7A7A1), and *B.subtilis* (P37528). The red boxes with white letters show identity, red letters show similarity in the region between amino acid groups, and black letters represent the non-conserved region, with TT being a strict β turn and TTT a strict α turn. The sequence alignment was prepared by Espript3 [105].

The orthologs of Pdx2 (YaaE, SNO, and PdxT) are reported and widely distributed at various evolutionary levels i.e. in archaea, bacteria, and eukarya [57, 90, 103]. However, significant differences have been observed between the plasmodial and bacterial Pdx2, and the most distinct one is the helix α6 and loop β5-β6 insertion in the plasmodial sequences which are not found in the bacterial homologs (**Figure 8c** sequence alignment). In the *P.vivax* Pdx2 sequence, the insertion includes amino acids (103–112) starting with loop and ends in small helix α6. Another insertion in plasmodial $Pdx2_{172–177}$ forms β-strand (β10) and extends to strand β11, which are the potential interface regions with Pdx1 counterpart [69, 103]. Further the *P.vivax* Pdx2 helix α6′ is a loop region in the reported *P.falciparum* Pdx2 (2ABW) [103] **(Figure 8b, c)**, possibly due to a single residue mutation (E166Q). Whereas in bacterial structures the residue at this position is missing, as shown in **figure 8c**.

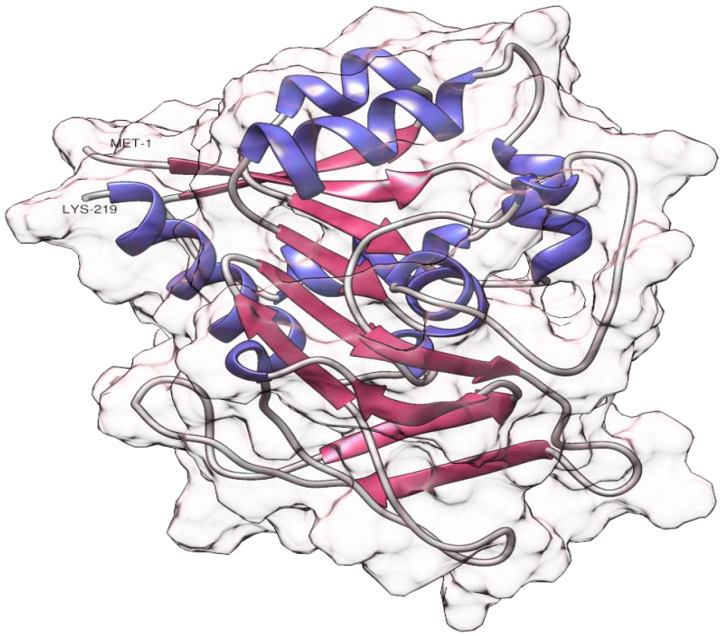


Figure-9. *P.vivax* Pdx2 3-dimensional modelled structure.
P.vivax Pdx2 3D structure is generated by I-Tasser [106] and prepared by UCSF Chimera [102] with secondary structure elements shown in different colors. N and C termini are labeled. Helices are shown in blue, strands in purple, and loops in light green color.

2.6.3. Coordination between Pdx1 and Pdx2

For product PLP formation, both Pdx1 and Pdx2 carry out the synthase and glutaminase reactions, respectively [68] in a series of reaction steps. Pdx2 binds to Pdx1 in a 1:1 ratio in which 12 Pdx2 subunits surround the dodecameric Pdx1 in a way stacked rings like studs, each interacting with one Pdx1 partner. Further structural investigations of the bacterial Pdx1-Pdx2 complex show that Pdx2 binds to the N-terminal side of Pdx1, with its active site present at the proteins interface [90, 97], as shown in **figure 10**.

Interestingly, in *in-vitro* investigations, it has been shown that Pdx1 exhibits the ability to detach from Pdx2 and can synthesize the product PLP utilizing NH_3 from ammonium sources like NH_2SO_4 or NH_4Cl [87]. This *in-vitro* NH_3 utilization has not been evidenced *in-vivo* until now, lacking an explanation of whether Pdx1 can also take NH_3 from the cellular environment. However, the *in-vitro* enzymatic investigations have shown that PLP synthesis happens at a very low rate without Pdx2 association to Pdx1, indicating Pdx2 attachment increases the PLP synthesis [87]. It suggests that there is domain communication between both counterparts for PLP synthesis. However, Pdx2 lacks glutaminase activity and is inactive in the absence of its counterpart Pdx1 [61].

Further studies which reported some preliminary domain-communication in Pdx1 and Pdx2, include the ordering of the helix αN from Pdx1 after attachment to Pdx2. The helix αN is spatially located at the interface of both proteins and is in close proximity to the Pdx2 active site pocket **(Figure 10).** It exhibits a key role in Pdx complex formation and Pdx2 activation [90, 94]. The Pdx1 N-terminus also form β-strand (βN) preceding helix αN adjacent to the Pdx2 β-strands architecture (antiparallel to β7 region) and establishing contacts to Pdx2. This phenomenon is reportedly called β-completion in bacterial GATases [94]. However, it has been reported that the plasmodial Pdx complex does not follow the β-completion phenomenon where βN orients towards Pdx1 and lacks establishing contact with Pdx2 [69]. However, this needs further investigations as the reported plasmodial Pdx complex is a chimeric complex structure with a loop instead of β-strand in the βN region. Further, in *P.vivax* Pdx1, the elongated N-terminus with extra six residues (MENAKR) **(Figure 10)** which are absent in the sequence of the reported homologs may answer this evolutionary alteration and needs to be explored in terms of the structural and functional investigations.

The functionally important oxyanion hole structure present in Pdx2 is involved in communication between both the proteins deserve a few lines. After binding to Pdx1, Pdx2 carries out glutamine hydrolysis, catalyzed by the catalytic triad. Glutamine hydrolysis is facilitated by the conserved oxyanion hole. Pdx1 Pdx2 complexation involves the re-organization and flipping of this oxyanion hole present in the loop region of β3-α3, in the proximity of the active site. This flip facilitates the transfer of ammonia, produced at the active site of Pdx2 [69, 90].

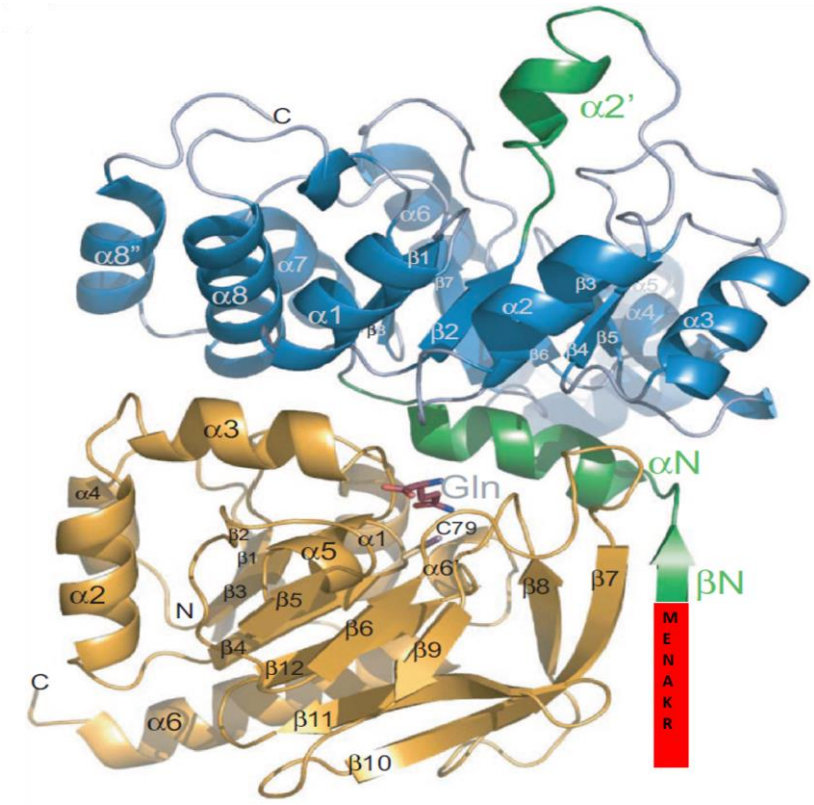


Figure-10. Ordering of the important structural elements of Pdx1 at complex formation. Pdx complex with Pdx1 shown in blue and Pdx2 in orange. Upon interaction and complex formation with Pdx2, three functionally significant structural elements ordered in Pdx1 (shown in green). These elements (βN, αN and α2′ helix), all highlighted in green, are present at the N-terminal of Pdx1. The β-strand βN is involved in β-completion with β-strands β7/β8 of Pdx2 subunit. Helix αN is present at the interface of both proteins and covers the Pdx2 active site. Glutamine (Pdx2 substrate) is shown in stick with catalytic triad residue Cys79 of *Bs*Pdx2. Figure is adapted from Marco Strohmeier *et al.* 2006, PNAS, vol. 103, no. 51, 19284–19289 [90].

2.7. Plasmodial PLP utilization or biosynthesis in RBCs: *An enigma*

Bacteria, plasmodium, fungi, and plants synthesize vitamin B_6 *de novo* via the pyridoxal 5-phosphate biosynthesis pathway, whereas humans acquire it through food and are entirely dependent on its uptake [107]. In Plasmodium the *pdx1* and *pdx2* genes, encoding Pdx1 and Pdx2 proteins respectively, are reported with their expression profile

in the intra-erythrocytic stages of Plasmodium indicate that both Pdx1 and Pdx2 proteins are localized in the cytosol. Additionally, the biochemical investigations revealed that Plasmodium possesses functional Pdx proteins, validating the presence of a functional vitamin B_6 biosynthesis pathway [108]. During the intra-erythrocytic life stages of Plasmodium, the parasite requires proper mechanisms of survival to cope with different host conditions. Since PLP is an essential cofactor involved in various enzymatic reactions and Plasmodium must acquire the continuous cellular PLP demands.

It is known that haemoglobin binds Pyridoxal (PL) and its phosphorylated form PLP (vitamin B_6) might serve as a PLP source for the parasite [109] as it utilizes more than 75% of the host haemoglobin [110]. Since erythrocytes are the major carriers of pyridoxal 5-phosphate (vitamin B_6) in mammals for storage and maintaining its balance. Plasmodium is also metabolically active and multiplies in RBCs during infection, highlighting the question "why Plasmodium utilizes its own energy for PLP production." One possible reason can be that vitamin B_6 is very tightly bound to albumin and haemoglobin in serum and RBCs, thus might not be accessible to the parasite [107]. To date, there is no convincing evidence, showing the erythrocytes vitamin B_6 contribution in Plasmodium nutrition. The second possible reason is that the vitamin B_6 vitamers in the form of pyridoxamine (PM) and pyridoxine (PN), and their phosphorylated vitamer forms (PNP, PMP) need to be converted into PL and/or PLP in the RBCs before delivered to the Plasmodium. The reason is, human homologs of PMP and PNP oxidases have not been discovered yet in the plasmodial genome, suggesting their conversion to PLP might not be possible in Plasmodium. This also requires RBCs to be more efficient in vitamin B_6 vitamers conversion. However, no such data is reported with an excess of PNP and PMP oxidases in RBCs at the time of active plasmodial infection.

In this thesis, I will discuss in detail Pdx1 crystal structures from *P.vivax* and *S.aureus*, solved after utilizing synchrotron radiation at DESY, Hamburg. Pdx complex from *P.vivax* has also been crystallized, but the diffraction data from the crystals was very weak and not promising. More to this, I have solved the solution structure (SAXS) models of Pdx1, to investigate its oligomeric nature, from both species utilizing P12 beamline at synchrotron source DESY, Hamburg. Moreover, solution scattering SAXS

data of Pdx2 and Pdx complex from both species (*P.vivax* and *S.aureus*) have also been analyzed.

2.8. Macromolecular X-ray crystallography and SAXS scattering

Macromolecular X-ray crystallography is a field where X-rays determine the accurate and precise structure models of biological molecules (proteins, DNA RNA) at atomic resolution. Macromolecular X-ray crystallography may also provide detailed molecular-level structural information of large biological assemblies. Significantly, it can also provide a detailed enzymatic mechanism as potential drug targets, leading to the structure-aided drug design. At a low-resolution level, small-angle X-ray scattering (SAXS) is applied for biological macromolecules, determining the overall shape and dispersity of macromolecules in solution.

SAXS aim to obtain low-resolution structural information about the size ranging from nanometres to a micrometer [111], overall shape, as well as the flexibility of the macromolecules randomly oriented in the solution. In a standard SAXS experiment, monodisperse macromolecule solutions are placed in a quartz capillary (bio-inert towards biological materials). The capillary with the sample is irradiated by an X-ray beam for about 50 milliseconds exposure time, with the sample continuously flowing through the capillary throughout the measurement. X-rays are scattered by the macromolecules present in the solution and the scattered X-rays in a scattering intensities profile are recorded by a photon-counting X-rays detector. To obtain the scattering intensities profile of the macromolecule in solution, the intensity of the buffer alone is recorded and then subtracted from the scattering pattern of the macromolecules in solution in automated data acquisition and processing steps [112]. From the radially integrated scattering profile, the size and overall solution shape of macromolecules can be derived [113]. The scheme of SAXS operation, processing, and application is illustrated in **figure 11.**

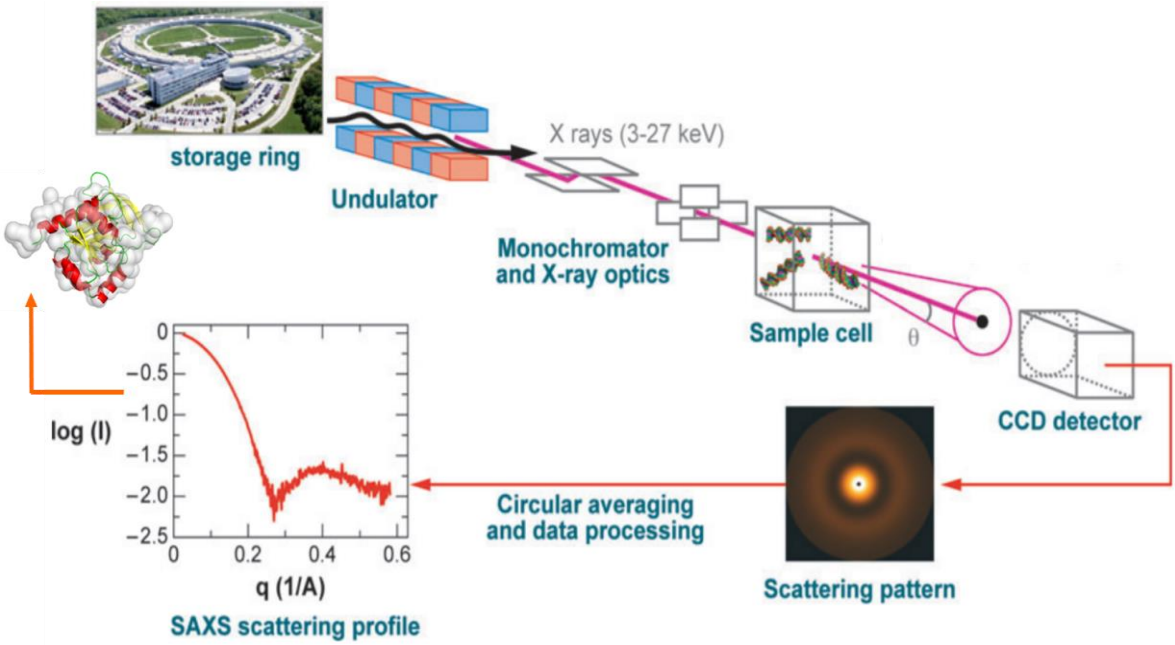


Figure-11. Schematic representation of a SAXS beamline at a synchrotron source.

The X-rays scattering intensity scattered from the sample is generally recorded as a function of momentum transfer S, ($S = 4\pi \sin(\theta)/\lambda$, where 2θ is the scattering angle and λ is the X-rays wavelength used for sample exposure). The collected scattering data have information about the size, mass, and shape of molecules, revealed after different softwares interpretation and calculation. The figure is adapted and modified from Lipfert Jan and Doniach Sebastian, Annu. Rev. Biophys. Biomol. Struct. 36, 307-327, 2007 [114].

2.8.1. Fundamentals of X-rays diffraction

X-rays are electromagnetic radiations with their wavelengths in the order of atomic distances i.e. 0.1 nm (1 Å). Whereas suitable X-rays used in the macromolecular crystallography are in the range between 0.5 Å and 1.6 Å wavelength. As of today, we enjoy more automation in protein purification, crystallization, and fine-tuneable robust synchrotron X-ray radiation sources. The pioneer macromolecular crystallographers and their efforts were not less than heroic in solving the macromolecular structures. The reason structural biology and crystallography produce several Nobel Laureates. With more advancements from manual to computer-aided structure solving, macromolecular X-ray crystallography improves a lot in the last few decades. As a result, over the past few years, X-ray crystallography has matured into a very rapid and automated method [115]. Once target crystallization conditions are optimized and established, the crystal screening procedure becomes then straight-forward with fewer efforts due to the availability of highly automated latest generation synchrotron sources [116]. Furthermore, the establishment and recent advancements in the photon counting

detector technology, and well-established user-friendly GUI (Graphical-User-Interface) data processing methods are a top-up. With the above-mentioned developments in technology, X-rays macromolecular structure determination now typically takes only a few minutes. It enables the screening of several hundred samples per day and with the automatically generated preliminary structure in the next few minutes.

To adequately investigate macromolecules via X-ray crystallography, the biological sample in a crystalline form is a pre-requisite, where the molecules are arranged orderly in three-dimensions, held together by noncovalent interactions [117]. A crystal (cryo-cooled and transferred from the crystallization drop) is typically mounted between the incoming beam from the X-ray source and an X-ray detector. The macromolecule/protein crystal sets to the narrow beam of X-rays originating from the source in the simplest form an X-ray tube or rotating anode to the complex particle storage rings producing synchrotron or FEL radiations (**Figure 12**). When protein crystals are exposed to the X-rays, they scatter X-rays in specific directions displaying a definite pattern called diffraction pattern. The diffraction pattern is imprinted on the X-ray detector, with the latest PILATUS photon-counting detectors. Diffraction is caused by the interaction of the X-rays with the atoms inside the crystals, to be more specific with the electrons of the atoms. The biomolecule e.g. protein unique diffraction pattern that emerges from the crystal contains valuable meaningful information about the structural arrangements of the biomolecules in the asymmetric unit and unit cell of crystals **(Figure 12)** [118].

For subsequent measurement of the directions and intensities of reflections, the crystallographers collect diffraction data from all the required orientations of the crystal in reference to the incoming X-ray beam. The collected diffracted data is a list of intensities for each diffraction point in the three-dimensional reciprocal lattice and from the intensities (which depends on the amount of electron density) of the obtained reflections, information about the shape of individual molecules in the crystal are obtained. However, for accurate shape determination, the significant phase of the beam (incoming waves) cannot be interpreted directly, because the X-ray detectors are sensitive only to the incoming energy and this energy is independent of the wave phase. If the reflection phases are known, then the electron density of the atoms in the

asymmetric unit and the unit cell can be obtained by applying the Fourier transformation.

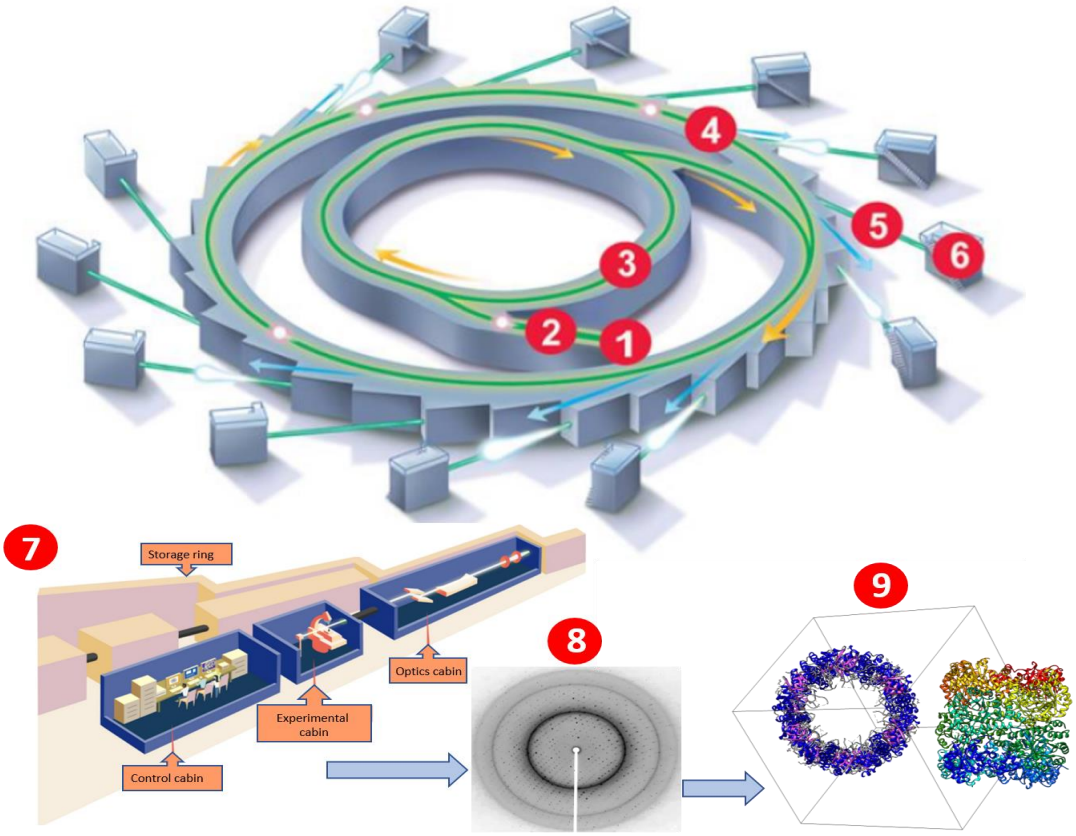


Figure-12. Synchrotron radiation source with main components.
(1) Electron gun which produces electrons (2) linear accelerator which accelerates the electrons to high speed (3) booster ring increases the energy of electrons (4) storage ring maintains the energy of the electrons which produce X-ray radiations (5) beamline setup for experiments (6) experimental hutch with one of its part is dedicated for beam size control (7) three main parts of a beamline for users access (8) an obtained diffraction pattern from a crystal, exposed to the incoming X-rays (9) the obtained crystal structure after diffraction data processing. The figure is updated from the DLS-CCP4 data collection and structure solution workshop 2016 [119].

2.8.2. Indexing and scaling of diffraction data

The reflection positions (*hkl*) obtained from diffraction spots are important in determining the phase information of the biomolecules. The position of a reflection is the angle by which the diffracted beam diverges from the X-ray beam with reference to the origin coordinates (*h*,*k*,*l* = 0,0,0). The indexing, integration, and scaling procedure consist of assigning the unit cell parameters based on the reflection positions and symmetry of the diffracted reflections. The automated softwares, like XDSAPP software [120], reads the reflection or spot positions providing information about the crystal lattice and calculate the unit cell parameters. After the unit-cell parameters are

obtained, then integration of the spots can be performed where all the raw pixel reflection intensities are converted into unique single reflections. The process consists of estimating the intensities of the reflected spots with an error value to each intensity after X-rays background subtraction. In the end, all values for the intensities are listed in the form of Miller indices (*hkl*) with respective intensities I(*hkl*) along with their background subtraction errors σI (*hkl*) for each observed reflection.

These integrated reflections containing many identical and symmetry-related reflections are then indexed and scaled in data reduction and scaling process utilizing computer-aided software like XDSAPP [120]. This generates unique reflection datasets corresponding to the asymmetric unit. The indicators of the quality of optimized and reduced data are then reflected in the signal-to-noise ratio ($I/\sigma I$), R-merge, and also redundancy in each shell and the entire dataset [117]. The statistical values and details for each dataset are presented in this thesis for Pdx1 *P.vivax* and Pdx1 *S.aureus* are provided in tables in the Results section.

2.8.3. Crystals unit cell and Asymmetric Unit (ASU) with protein molecules arrangement

A crystal is a whole lattice (array) of its unit cells stacked end to end in three-dimensional space. Whereas the simplest repeating unit in a crystal is called a unit cell. The unit cell can also be defined in terms of the lattice points i.e. the points in space about which the crystallized particles are free to vibrate in a crystal. Atoms/particles in the crystal are arranged in real space. To describe the diffraction phenomena and obtain further meaningful information, the non-physical reciprocal space concept is utilized. Since, the reciprocal-lattice and lattice spacings are simply the inverse (Fourier transform) of the real-lattice and lattice spacings, connecting diffraction spots and the distances between the reflections in the diffraction pattern imprinted on the detector. Therefore, it is possible to measure the unit cell dimensions from the position of the diffraction spots (the angles of reflections) and the distance between the reflection spacings. Further, the crystal unit cell dimensions are designated by six lattice parameters: the lengths parameters (a, b, c) of the three unique axes **a, b, and c**; and the three unique angle parameters α, β, and γ [118].

What crystallographers deposit in the various crystallographic databases like protein data bank (**PDB**) are the coordinates of the atoms in the asymmetric unit. The

asymmetric unit of a specific space group is that part of the crystal unit cell which can be efficiently utilized to build the complete unit cell by applying symmetry of the space group. ASU can also be defined as the largest accumulation of biomolecules in the unit cell possessing no symmetry elements but can be juxtaposed to other identical unit cell components by applying symmetry operations. For instance, in proteins in the most simple case, the asymmetric unit is a single functional protein molecule in the unit cell [118].

Since protein molecules are asymmetric intrinsically, composed of amino acids (that are chiral in nature) which create larger chiral secondary structure elements like right-handed helices or beta-sheets. In a most simple case, if only one protein molecule occupies the asymmetric unit as well as the unit cell, then the cell itself is chiral having no symmetry elements. This situation is rare in protein crystallization, and in most cases, the unit cell contains several identical molecules or oligomeric units in an arrangement possessing symmetry elements. Because of the intrinsic chiral nature of proteins, the symmetry operations and elements needed to describe protein crystals ASU and unit cells include only translations, rotations, and screw axes (which are combined rotations and translations). This limitation on the symmetry of unit cells containing asymmetrical objects reduces the number of space groups for chiral molecules i.e. proteins from 230 to 65 [118].

Further, protein molecules show a rather constant density of 1.35 gcm^{-3} or in other words 1.35 $Å^3$/dalton, independent of the nature of the protein. But because of the solvent content, crystallized proteins do not follow this rule of displaying a constant density. Often half or more of the crystal unit cell volume, called Matthews Volume, is occupied by solvent molecules and on average protein molecules in crystals occupy a volume of 2.4 $Å^3$/dalton with a varying range from 1.9 to 4.2 $Å^3$/dalton. This “Matthews Volume” is very important and useful in terms of finding the biological monomeric or oligomeric functional unit of proteins [121]. The structures solved in this thesis are multimeric or oligomeric (octamer and dodecamer) indicating the significance of proteins molecules multimeric functional units arrangement in ASU and the unit cell. Pdx1 *P.vivax* ASU consists of two space group symmetries with each in dodecamer form. Whereas Pdx1 *S.aureus* with 18 octameric space group symmetries in the unit cell complicate the information obtained for the oligomeric functional unit and structure

determination. It also makes it difficult to assign molecules arrangement in the three dimensions. Details of both structures are presented in the results section of the thesis.

2.9. AIMS of the Thesis

A robust resistance has been observed practically in human pathogenic plasmodial and bacterial *S.aureus* species against almost all of today's available drugs. This highlights the necessity of novel antimalarials and/or antibiotics which should be effective against the novel metabolic pathways of the microbes, particularly those absent in humans, like the novel vitamin B_6 synthesis pathway. Vitamin B_6 biosynthesis pathway is reported to be present at different evolutionary levels reported from prokaryotes to eukaryotic plants, whereas absent in humans. Further, the involvement of its product (PLP) in various essential enzymatic pathways as a cofactor, makes the PLP biosynthesis pathway a very interesting drug target. PLP biosynthesis requires three substrates i.e. two sugars ribose 5-phosphate and glyceraldehyde 3-phosphate, whereas the third substrate $_L$-glutamine is utilized by the Pdx2 subunit producing neutral ammonia. The produced ammonia moves across the intermolecular tunnel at the Pdx1-Pdx2 interface and is finally utilized at the active site of the Pdx1 subunit.

To date, the literature lacks the detailed comparative bioanalytical and structural investigations reported for either of the subunit and or their complex. Current data dispense information mainly describing the Pdx2 mode of activation. Further, the eukaryotic plasmodial Pdx1, Pdx2, and Pdx complex are reported to form fibers, induces oligomerization, and hence are considered difficult targets for structural investigations. So far, no or limited drugs designing investigative reports are available for the PLP synthase proteins.

Therefore, concerning the scarcity of the compulsory literature in the PLP project, this thesis was aimed at the detailed biophysical and bioanalytical investigations of the *Plasmodium vivax* Pdx proteins. The thesis was also aimed at the comparative biophysical and structural evaluation of the *P.vivax* and *S.aureus* Pdx complex and study the dynamicity of its 24meric assembly by time-resolved X-rays investigations. Subsequently, examine the interface regions of the Pdx1-Pdx2 proteins for future drug discoveries.

The Pdx1 subunit, forming the dodecamer, constitutes the main essential core of the PLP synthase complex as observed in some bacterial 3D crystalline structures. However, in unicellular eukaryote (yeast) Pdx1 is observed in hexamer form in both solutions, as well as in the crystal structure with differences in the enzymatic activity obtained from biochemical data when compared to the available prokaryotic data. Whereas, eukaryotic plasmodial Pdx1 is observed in the dodecameric form with differences in the structural elements, important functionally. Based on these structural and functional differences, the thesis was also aimed at the comparative structural and functional investigations of the Pdx1 proteins from *Plasmodium vivax* and *Staphylococcus aureus*.

The main motivation behind my PhD investigations was to identify the structurally important elements of Pdx1 subunits by comparing the crystal and solution structures.

3.0. Materials and Methods

3.1. Materials

3.1.1. Chemicals

All chemicals utilized during my PhD studies were purchased in analytical grade from the companies Sigma-Aldrich GmbH (Germany), (Sigma, USA), AppliChem GmbH (Germany), Hampton Research (USA), Merck KGaA (Germany), Qiagen (Germany), (Biomol, Germany), Thermo Fisher Scientific Inc. (USA), Carl Roth GmbH (Germany), Serva Electrophoresis GmbH (Germany), Molecular Dimensions Limited (UK), Cayman Chemical Michigan (USA), Jena Bioscience (Germany).

3.1.2. Consumables

Falcon tubes, reaction tubes, serological pipettes, and pipette tips utilized in PhD research were purchased from the company Sarstedt (Germany). Further consumables are listed below in table 1.

Table 1: Consumables utilized during PhD work

Description	**Manufacturer/Supplier**
96-well Vapor batch diffusion plates	Douglas Instruments Limited, UK
48-well MRC maxi plates	Jena Bioscience GmbH, Germany
96-well SWISSCI MRC2 plates	Hampton Research Corp., USA
Dialysis tubing, MWCO 6-8, 12-14 kDa	REPLIGEN, Germany
24-well Linbro plates	Jena Bioscience GmbH, Germany
Amicon Ultra 0.5, Ultra 05, Ultra 15, of MWCO 3,10 and 30 kDa	Merck Millipore (Merck KGaA), Germany
Cellulose acetate membrane filter 0.2 µM	VWR International LLC, USA
Filters for PEG filtration (0.45 um)	VWR International LLC, USA
Coverslips	VWR International LLC, USA
Econo-chromatography columns 100 cm, 20 cm	Bio-Rad (Germany)
Highly viscous vacuum grease	GE Bayer Silicones
Protein marker	Thermo Fisher Scientific Inc., USA
Ni-NTA agarose matrix	Qiagen, Germany
Syringes 1 mL, 2 mL, 5 mL, 10 mL, 20 mL, 50mL	VWR International LLC, USA
Standard calibration protein kit	Merck Millipore (Merck KGaA), Germany
ZipTip®C18 Pipette tips	Merck Millipore (Merck KGaA), Germany

3.1.3. Laboratory Equipment

The following table lists the laboratory equipment I have utilized during my PhD work (Table 2).

Table 2: Laboratory equipment used in PhD work

Device	**Instrument**	**Manufacturer/Supplier**
CD spectrometer	J-815 CD	Jasco Inc., USA
Centrifuges	5415R/ 5415C/ 5804R/	Eppendorf AG, Germany
	5810R MinispinPlus	Eppendorf AG, Germany
	Multifuge X3R	Thermo Fisher Scientific Inc., USA
Crystallization Robots	Honeybee 961	Genomic Solutions, UK
	Oryx 4	Douglas Instruments Limited, UK
Crystal plate incubator 4 °C and 20 °C	RUMED 3001 incubator	Rubarth Apparate GmbH, Germany
UV light source	Crystal Score device	Diversified Scientific Inc, USA
DLS device	SpectroLight 300	Xtal Concepts GmbH, Germany
	SpectroLight 600	Xtal Concepts GmbH, Germany
FPLC system	ÄKTA Prime with fraction collector and UV detector	GE Healthcare, UK
	ÄKTA Purifier P-901 with fraction collector and UV detector	GE Healthcare, UK
Freezer -20°C	Liebherr premium	Liebherr International S.A., Switzerland
Freezer -80°C	B35-85	FRYKA-Kältetechnik GmbH, Germany
Agarose gel electrophoresis chamber	PAGE chamber	Serva Electrophoresis GmbH, Germany
	Perfect Blue Gel system Mini M	PEQLAB Biotechnologie GmbH, Germany
Chromatography columns	Hi Load 16/60 Superdex 75, Hi Load 26/60 Superdex 200, and Superose 6 increase 10/300	GE Healthcare, Uppsala, Sweden
Imaging	Crystal Score	Diversified Scientific Inc., USA
	Leica TCS SP8 Confocal Platform	Leica Microsystems GmbH, Germany
Incubator 37°C and 60°C	BD 56	BINDER GmbH, Germany
Incubation shaker	KS 3000 I control	IKA Werke GmbH & Co. KG, Germany
	Innova 4330	New Brunswick Scientific Co., Inc. (Eppendorf AG), Germany
Magnetic stirrer	VMS-A	VWR International LLC, USA
Mass Photometer	Refeyn OneMP mass photometer	Refeyn Ltd. UK
Microbalance	Sartorius TE3102S CP2245-OCE	Sartorius AG, Germany
pH Meter	Seven Easy	Mettler-Toledo Inc., USA

Device	Instrument	Manufacturer/Supplier
Spectrophotometers	Nanodrop 2000c	Thermo Fisher Scientific Inc., USA
	Nanodrop Lite	Thermo Fisher Scientific Inc., USA
Pipettes	Eppendorf Research, 2.0 µL, 10 µL, 100 µL, 1.0 mL	Eppendorf AG, Germany
Sonifier	Branson sonifier 250	Emerson Electric Co, USA
Negative staining glow discharger for EM grids	GloQube glow discharger	East Sussex, UK
SDS-PAGE	Four Gel Caster (SE275)	Hoefer Inc., USA
	SE260 Mighty Small II Deluxe Mini electro. unit	Hoefer Inc., USA
	EV 231 (Power supply)	PEQLAB GmbH, Germany
Thermomixer	Comfort	Eppendorf AG, Germany
Vortex mixer	VF2	IKA Werke GmbH & Co. KG, Germany
X-ray source	IµS	Incoatec GmbH, Germany
Detector	-	Mar Research GmbH, Germany

3.1.4. Beamlines attended during PhD studies

During my PhD studies, I had the opportunity to attend beamtimes at the beamlines located at the DESY synchrotron radiation source mentioned in table 3 for SAXS solution scattering and crystal diffraction data collection.

Table 3: Beamlines attended at synchrotron source DESY

S.No.	Beamline station	Source and affiliation	Purpose of attendance
1	P11	DESY, PETRA III, DESY, Hamburg, Germany	Diffraction data collection
2	P12	EMBL, PETRA III, DESY, Hamburg, Germany	SAXS solution scattering data collection
3	P13	EMBL, PETRA III, DESY, Hamburg, Germany	Diffraction data collection

3.1.5. Media, Antibiotics and Buffers solutions

The media, antibiotics, and buffer solutions were prepared in double-distilled water (ddH_2O). The pH of media and buffers was adjusted with solutions of HCl or NaOH for every buffer where needed. The media and antibiotics and buffer solutions are presented in Tables 4 and 5 respectively.

Table 4: Growth media and antibiotics used in bacterial cultures preparation.

Growth media used	**Growth media composition**	
Luria Bertani, Lennox (LB) media	Tryptone (10 $g \cdot L^{-1}$), Yeast extract (5 $g \cdot L^{-1}$), NaCl (10 $g \cdot L^{-1}$)	
LB-Agar media	1.5 % (w/v) of agar media added in LB	
Antibiotic	**Stock solution**	**Working concentration**
Ampicillin	25 $mg \cdot mL^{-1}$ in deionized H_2O	100 $\mu g \cdot mL^{-1}$
Chloramphenicol	34 $mg \cdot mL^{-1}$ in 100% ethanol	34 $\mu g \cdot mL^{-1}$
Carbenicillin	25 $mg \cdot mL^{-1}$ in deionized H_2O	100 $\mu g \cdot mL^{-1}$

Table 5: Generally used buffers and solutions

Name of the buffer/solution	**Buffer/solution Composition**	**Concentration**
Isopropyl-β-D-Thiogalacto-pyranoside (IPTG)	IPTG	1 M
Phenylmethylsulfonylfluoride (PMSF)	PMSF	100 mM (dissolved in 50% isopropanol)
Coomassie staining solution	Acetic acid	10% (v/v)
	Isopropanol	25% (v/v)
	Coomassie (Brilliant Blue R-250)	0.1% (w/v)
Coomassie destaining solution	Acetic acid	20% (v/v)
2 x sample buffer SDS PAGE	Tris-HCl (pH = 6.8)	20 mM
	SDS solution	4% (*w/v*)
	Glycerol	20% (*v/v*)
	Dithiothreitol (DTT)	50 mM
	Bromophenol blue	0.04% (*w/v*)
5 x native sample buffer native PAGE	Tris-HCl (pH = 6.8)	50 mM
	Glycerol	50% (*v/v*)
	Bromophenol blue (BPB)	0.1% (*w/v*)
Separating gel buffer	Tris-HCl (pH = 8.8)	1.5 M
	SDS solution	0.4% (w/v)
Stacking gel buffer	Tris-HCl (pH = 6.8)	0.5 M
	SDS solution	0.4% (w/v)
10 x electrode buffer	Glycine	1.920 M
	Tris-HCl	250 mM
	SDS solution	1% (w/v)
1 x Native PAGE electrode buffer	Glycine	192 mM
	Tris-HCl	25 mM

3.1.6. Buffer utilized for Pdx proteins purification

3.1.6.1. Buffers used for *P.vivax* and *S.aureus* Pdx proteins purification

All the buffers used for the purification of Pdx proteins (Pdx1, Pdx2, and Pdx complex) from *P.vivax* and *S.aureus* are listed in table 6.

Table 6: Ni-NTA and SEC buffers used for *P.vivax* and *S.aureus* Pdx proteins purification

Name of buffer/solution	Composition	Working concentration
Lysis buffer	NaH_2PO_4 (pH = 8.0)	50 mM
	NaCl	300 mM
	MTG	2.0 mM
	EDTA	1.0 mM
	Imidazole	10.0 mM
Washing buffer	NaH_2PO_4 (pH = 8.0)	50 mM
	NaCl	300 mM
	MTG	2.0 mM
	EDTA	1.0 mM
	Imidazole	50 mM
Elution buffer	NaH_2PO_4 (pH = 8.0)	50 mM
	NaCl	300 mM
	MTG	2.0 mM
	EDTA	1.0 mM
	Imidazole	250 mM
Pdx1 SEC buffer	Tris–HCl (pH = 8.0)	20 mM
	NaCl	200 mM
	MTG	2.0 mM
	EDTA	1.0 mM
	Glycerol	2 %
Pdx2 SEC buffer	HEPES (pH = 7.0)	20 mM
	KCl	200 mM
	MTG	2.0 mM
	EDTA	1.0 mM
	Glycerol	3.0%
Pdx complex SEC buffer	Tris–HCl (pH = 8.0)	20 mM
	NaCl	200 mM
	MTG	2.0 mM
	EDTA	1.0 mM
	Glycerol	2.0%
	L-glutamine	20 mM

3.1.7. Buffers utilized in SDS-PAGE

Buffers used in SDS-PAGE include buffers for gel casting and SDS-PAGE preparation buffers, presented in tables 7 and 8.

Table 7: Buffers and solutions used for SDS-PAGE gel casting

Gel type	**Component**	**Volume**
Stacking gel (4%)	acrylamide/bisacrylamide (37.5:1): 30% solution	2.0 mL
	Stacking gel buffer	3.8 mL
	10% (w/v) SDS solution	150 µL
	TEMED	15 µL
	10% (w/v) APS solution	75 µL
	ddH_2O	9.2 mL
Separating gel (12%)	acrylamide/bisacrylamide (37.5:1): 30% solution	12.0 mL
	Stacking gel buffer	7.5 mL
	10% (w/v) SDS solution	300 µL
	TEMED	15 µL
	10% (w/v) APS solution	150 µL
	ddH_2O	10.2 mL

Table 8: SDS-PAGE preparation buffers

Solution	**Composition**
sample buffer: 2 x	10% SDS (w/v) solution: 2.0 mL, 0.5 M Tris-HCl 1.25 mL pH = 6.8, 100% glycerol: 2.5 mL, 0.5% Bromophenol Blue (w/v): 0.2 mL, 0.05% DTT in 3.55 mL. All these solutions were prepared in ddH_2O.
APS solution	10% in ddH_2O
Electrode buffer: 10 x	Tris: 30.0 g, Glycine: 144 g, SDS: 10.0 g, all in 1L of ddH_2O
Separating gel buffer	1.5 M Tris/HCl, pH = 8.8 in 100 mL of ddH_2O
Stacking gel buffer	0.5 M Tris/HCl, pH 6.8 in 100 mL of ddH_2O
SDS buffer	10% (w/v) in ddH_2O
Staining solution	0.25% (w/v) Coomassie (Brilliant Blue G-250), 25% (v/v) Isopropanol and 10% (v/v) acetic acid.

3.1.8. Stock solutions of buffers and crystallization screens utilized for Pdx proteins crystallization

Table 9: Stock solutions of buffers prepared to crystallize Pdx proteins

Stock solution	Concentration
LiCl	1 M
	2 M
Lithium citrate	2 M
Lithium sulfate	2 M
Ammonium sulfate	2 M
$MgCl_2$	2 M
Ethanol	1 M
Sodium malonate	50 %
Tris	1 M

Table 10: List of crystallization screens used in Pdx proteins crystallization

S.No.	Name of Screen	Supplier
1	$AmSO_4$-Suite	QIAGEN N.V., Netherlands
2	ComPAS-Suite	QIAGEN N.V., Netherlands
3	JCSG-Suite	Molecular Dimensions Ltd., USA
4	MORPHEUS Complete HT-96	Molecular Dimensions Ltd., USA
5	PACT premier	Molecular Dimensions Ltd., USA
6	Structure	Molecular Dimensions Ltd., USA
7	Stura	Molecular Dimensions Ltd., USA

3.1.9. Bacterial strains utilized

Different bacterial (*Escherichia coli*) strains used for Pdx genes plasmid transformations and recombinant gene expression are presented in table 11.

Table 11: Bacterial strains used for transformation and expression of Pdx genes

Bacterial Strain	Genotype	Supplier
BL21 (DE3)	F- *ompT hsdS*B (rB^- mB^-) *gal dcm rne131*(DE3)	Invitrogen (Thermo Fisher Scientific Inc.), USA
BL21 (DE3) Rosetta	F- ompT hsdSB (rB^- mB^-) gal dcm (DE3) pRARE (CamR)	Novagen, Merck KGaA, Germany
BL21-CodonPlus-(DE3)-RIL	F- ompT hsdS(rB^- mB^-) gal dcm^+ Tetr gal λ [argU ileY leuW Cam^R]	Invitrogen (Thermo Fisher Scientific Inc.), USA
XL10-Gold	Tetrdelta- (*mcrA)183 delta- (mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F´ *proAB lacIqZDM15* Tn*10* (Tetr) Amy Camr]	Agilent Technologies, US

3.2. Molecular Biology

3.2.1. Preparation of chemically competent BL21-CodonPlus (DE3)-RIL cells

To transfer the plasmids DNA including the pdx genes, for subsequent Pdx proteins recombinant expression, different *E.coli* bacterial cells were prepared for chemical competency. For preparing the bacterial pre-culture, a single bacterial colony from a culture plate or some cells from a cold stored glycerol stock (200 µL) were inoculated with 10 mL of lysogeny broth (LB) media (with antibiotic chloramphenicol). Next, the culture media was incubated for about 8-10 hours at optimal 37°C temperature in a temperature-controlled shaker at a shaking speed of 180 rpm. The next day, the precultured cells were centrifuged at high speed, and the pellet was resuspended and transferred to 100 mL of LB media (with antibiotic chloramphenicol). The culture was incubated at 37°C temperature until the optical density at 600 nm (OD_{600}) reaches 0.6-0.7 in few hours. Afterward, the culture was cooled on ice for 15 min before spinning down at 4°C applying centrifugation at maximum speed. Once more, the supernatant was discarded, and the cell pellets were washed and resuspended in 5 mL of sterile 100 mM $CaCl_2$ solution. After another 15 min centrifugation step, with supernatant discarded, and the cells were resuspended in 5 mL of sterile 100 mM $CaCl_2$ solution supplemented with 10% glycerol. Aliquots of 50 µL were transferred to the reaction tubes and flash-frozen in liquid nitrogen and subsequent storage at -80°C for future utilization.

3.2.2. Transformation of the (DE3)-RIL cells with pdx genes plasmid

One aliquot of fresh frozen chemically competent (DE3)-RIL *E.coli* cells was thawed on ice, before 50-100 ng of plasmid DNA of *pdx1* and *pdx2* (in separate experiments) were pipetted into the reaction tubes.

Both *pdx1* and *pdx2* genes containing plasmids were obtained from BioCat (GmbH, Heidelberg, Germany), inserted in the pET31b(+) expression vector. The plasmid contains an ampicillin resistance gene for mutant selection and the lacI repressor for induction of the target gene expression with IPTG.

Chemically competent (DE3)-RIL cells with chloramphenicol resistance and plasmid DNA with ampicillin resistance were thawed on ice for 5 minutes till the frozen solution changes to a dense liquid. Cells were mixed with 2 µL of the plasmid DNA and incubated on ice for 20-30 min. In the following step, cells were heat-shocked for 30

seconds at 42°C and transferred onto the ice for 2 min. SOC media of 300 µL was added into the heat-shocked bacterial cells and incubated in the next step at 37°C temperature and 300 rpm speed for one hour, with mixing in between. Incubated cells were centrifuged for 30 seconds at maximum speed and the supernatant was discarded. The remaining pellet was resuspended in 200 µL of SOC media and plated on LB agar plates. The agar plate media contain both antibiotics, ampicillin and chloramphenicol. All this transformation protocol was performed in a sterile environment. The cultured plates after transformed bacteria streaking were incubated over-night at 37°C for microbial growth.

3.2.3. (DE3)-RIL (containing pdx genes) glycerol stock preparation

From Pdx plasmid transformed colonies grown on cultured petri plates, a single colony from the plate was streaked and grown in LB-media until it reaches the growth log-phase with OD_{600} of 0.6 to 0.8. The culture was preserved by the addition of 50% (v/v) glycerol to the culture. Stock aliquots of 500 µL were flash-frozen in liquid nitrogen and stored at -80°C for future utilization.

3.2.4. Isolation and purification of plasmids

The XL10-Gold strain resistant to tetracycline and chloramphenicol was used for the isolation and preparation of the plasmids DNA. XL10-Gold competent cells were transformed with the plasmid containing the pdx gene applying the same protocol as elaborated in methods section 3.2.2 and plated on the agar plate. The next day, DNA plasmid amplification was performed by incubating a colony from the plate in 5mL LB culture media at 37°C overnight. The amplified plasmid DNA was isolated from the cells using the DNA extraction kit i.e. peqGOLD mini-prep kit1 (PEQLAB GmbH, Germany), from the centrifuged pellet of the over-night grown culture following the manufacturer's instructions. The DNA concentration was determined using a Nanodrop spectrophotometer (Thermo Fisher Scientific Inc., USA) at a wavelength of 260 nm. The purity of the freshly isolated DNA was determined by calculating the absorption ratio at 260 nm and 280 nm. This value should be around 1.8 for pure DNA. Further, the eluted plasmid DNA containing pdx genes were stored at -20 °C.

3.3. Biochemical Methods

3.3.1. Recombinant expression of the pdx genes from *P.vivax* and *S.aureus*

Since crystallization demands a high amount of protein, therefore Pdx proteins were produced on large scale by recombinantly expressing the genes encoding the target Pdx proteins in BL21-CodonPlus-(DE3)-RIL cells, containing the pdx genes on a transformed DNA plasmid.

For large-scale protein production one colony of the (DE3)-RIL strain carrying the transformed plasmid, was added in 100 mL of LB medium with pre-added antibiotics ampicillin and chloramphenicol (Roth, Germany). It was further supplemented with 1.0 % glucose (to stop the leaky expression of the desired protein) overnight at 37°C, at speed of 180 rpm. For the main culture, 50 mL of the centrifuged and resuspended pellet preculture was transferred into 1 L of LB with the respective antibiotics inclusion. The *pdx1* and *pdx2* genes containing bacterial cultures were incubated separately at 37°C and 180 rpm speed until the OD_{600} value reached between 0.6 and 0.8, followed by 0.5 mM ITPG addition to each 5 L flask. Before IPTG induction, 1 mL from each culture was transferred to a reaction tube and saved for SDS-PAGE analysis. After 1 h, 2 h, 3 h, 4 h, and 5 h of post-induction incubation the OD_{600} was measured and 1 mL of culture was withdrawn and transferred to a reaction tube at each step. Every hour transferred and saved samples were centrifuged for a few seconds at maximum speed. Subsequently, pellets were resuspended in 50 µL of 2x SDS-PAGE buffer before and heated at 96°C for about seven minutes. Samples were applied on SDS-PAGE to analyze the optimal expression of Pdx proteins.

After 5 h of IPTG induction at 37°C, the expressed culture was transferred to centrifuging buckets for pellet production and centrifuged for 30 minutes at 8°C at a maximum speed of 4000 rpm. The supernatant from the centrifuged sample was discarded. Next, the cell pellets were resuspended in the PBS buffer and centrifuged again in 50 mL falcon tubes for about 20 min, at the same speed and temperature. The supernatant was then carefully discarded, and the remaining pellets containing the expressed Pdx proteins were stored at -20°C for future purification experiments.

3.3.2. Preparation of cleared cell lysates

The (DE3)-RIL cell pellets were thawed by letting them at room temperature and resuspended in 15 mL of lysis buffer. A spatula tip of lysozyme was added in powder,

and the cell suspension was placed on ice for about 20 min, followed by the addition of 100 µM PMSF protease inhibitor (Sigma, USA). The cells were sonicated on ice for 10 seconds with 30 seconds time intervals using a Branson sonifier 250 (Emerson Electric Co, USA). After clear homogenous suspension with color change (an indication of optimal sonication), the cell suspension was allowed to centrifuge for 30 min, at 8°C and maximum speed. The centrifuged pellet containing debris was separated from the soluble cell lysate (supernatant) comprising the desired proteins. The supernatant was filtered through a 0.2 µm filter before going to further purification steps. Small amounts of cell lysate and pellet were stored for SDS evaluation to confirm the optimal condition of sonication and protein solubilization.

3.3.3. His-tagged Pdx1 and Pdx2 proteins Ni-NTA affinity purification

After cell disruption and the supernatant containing Pdx proteins separation was applied on a previously equilibrated Ni-NTA agarose matrix (Qiagen, Germany) with lysis buffer containing 10 mM imidazole. The column was allowed to incubate on a roller rocker at 8°C for about 1 h. The flow-through was eluted, collected, and stored for SDS-PAGE to confirm the protein's binding efficacy to the affinity matrix. In the following step, the matrix was washed with four column volumes (CVs) of washing buffer (containing 50 mM imidazole) to remove any unbound proteins. The eluted fractions were stored for SDS-PAGE evaluation to confirm the binding affinity Pdx proteins to the affinity matrix. In the next elution step, the column was allowed to incubate with two CVs of the elution buffer containing 250 mM imidazole. The eluted Pdx1 and Pdx2 protein fractions were collected in separate reaction tubes and also stored for SDS-PAGE evaluation. The eluted Pdx proteins concentration in fractions collected in separate reaction tubes were determined by Nanodrop 2000c (Thermo Fisher, USA) at 280 nm.

3.3.4. TEV protease digestion and separation of cleaved Pdx proteins

TEV (tobacco etch virus) protease was used to remove the C-terminal His_6-tag from Pdx proteins at the TEV cleavage site cloned with the His-tag in the plasmid. The TEV cleavage was performed by overnight incubation of the Pdx proteins with the SEC purified TEV protease at a ratio of 1:50. Digestion was carried out in parallel at room temperature for 5 h and overnight at 8°C. Digested samples were passed through the Ni-NTA column and this time flow-through containing the Pdx proteins was collected and

analyzed by SDS-PAGE to determine and optimize the cleavage protocol. Less amount of *P.vivax* Pdx2 after digestion was indicative of its instability during cleavage and in the post-cleavage process. Therefore, the cleaved Pdx2 was dialyzed with a stabilizing buffer (SEC buffer containing a high amount of glycerol and DTT) before TEV removal. After dialysis of the cleaved Pdx2 protein, the mixture of the cleaved His-tag Pdx2 and TEV protease were allowed to incubate with the Ni-NTA resin. In the next step, the digested Pdx2 was collected in flow-through and stored in diluted form in reaction tubes.

3.3.5. Protein quantification by absorption measurements

The concentrations of the eluted Pdx protein solutions were measured photometrically. The absorption maximum of proteins is at 280 nm, the wavelength at which the aromatic amino acids, mainly tyrosine, tryptophan, and phenylalanine have maximum absorption. Further, the protein sequence can be used to estimate very accurately their extinction coefficient, important in estimating the molar concentration of the solution from its absorbance value. For *P.vivax* Pdx1 with a monomer of molecular weight 33000 Da, the extinction coefficient is 15470 cm^{-1} M^{-1} and for Pdx2 with a monomeric molecular weight of 25200 Da, its value is 15390 cm^{-1} M^{-1}. For *S.aureus* Pdx1 with molecular weight 32000 Da, is 12950 cm^{-1} M^{-1} and for *Sa*Pdx2 the extinction coefficient value is 7450 cm^{-1} M^{-1} for its molecular weight of 21000 Da. All the extinction coefficient values indicate the reduced states of the proteins and are calculated by the Prot-param tool on the Expasy server [122].

After estimating the molar extinction coefficient values, the concentration of Pdx proteins was determined by a nanodrop 2000c (Thermo Fisher Scientific Inc., USA), applying 1 or 2 μL of the sample. For measurement, 1 or 2 μL of protein buffer was applied and measurement was blanked two times. For protein concentration determination the procedure was repeated by applying 1 or 2 μL of protein solution. The molar extinction coefficient (ε) and the molecular weight values for the respective proteins were entered during each measurement, and the protein concentration was displayed by the software, with calculations based on the Lambert-Beer law, shown in **Eqn-1**.

$$A_{280} = \varepsilon c l \quad (1)$$

A_{280} is absorption at 280 nm
ε is molar extinction coefficient [M^{-1} cm^{-1}]
l is path length [cm]
c is protein concentration [mg/ml]

3.3.6. Regeneration of the Ni-NTA matrix

The Ni-NTA resin, with an affinity for His-tagged proteins, loses binding affinity with subsequent uses, and Ni-degrades from the resin and matrix regeneration is a multistep established protocol. For regeneration, the ofttimes used Ni-NTA matrix (with colorless or light greyish color) was first washed with water and then applying two (column volumes) CVs of the regeneration buffer (6 M Gu-HCl and 0.2 M acetic acid). Next five CVs of water, and afterward with three CVs of 2% (w/v) sodium dodecyl sulfate (SDS) solution. An ethanol (Roth, Germany) gradient from 25% (1 CV per washing step) to 100% (5 CVs) and back to 25% was applied. For complete removal of the ethanol, the resin was washed with one CV of water. To complex Ni^{2+} ions, the matrix was incubated with five CVs of 100 mM ethylenediamine-tetra-acetic acid (EDTA), pH 8.0 (Biomol, Germany), and was subsequently washed with two CVs of water. The matrix was then recharged for 10 min with two CVs of 100 mM $NiSO_4$ (AppliChem, Germany) containing buffer, washed with two CVs of water, two CVs of regeneration buffer, and one CV of water. In the last step, the matrix was equilibrated with protein lysis buffer or stored in 20% ethanol at 8°C. Details of different stock solutions utilized in Ni-NTA column regeneration are given in table 12.

Table 12: Ni-NTA agarose column regeneration solutions

Stock Solution	**Composition**
SDS solution	2% (w/v) in ddH_2O
EDTA solution	100 mM, pH 8.0 in ddH_2O
Regeneration buffer	6 M guanidine chloride, 0.1 M acetic acid
$NiSO_4$ solution	100 mM in ddH_2O
Ethanol solution	25, 50, 75% ethanol in ddH_2O, and 100%

3.3.7. Dialysis

Dialysis is a passive diffusion technique used to get rid of unwanted and leftover materials based on the molecular mass. It is also used to introduce/remove small

compounds through a semi-permeable membrane of a suitable molecular weight cut-off (MWCO) for the target proteins. This method generally works by either utilizing a dialysis membrane with a suitable MWCO or using an ultra-filtration membrane of different pore sizes for different molecular weight cut-offs. Dialysis membrane or ultra-filtration membrane (of 3 kDa/ 10 kDa/ 30 kDa MWCO) were used for routine dialysis. They were utilized either to modify the buffer composition of the Pdx proteins in between two subsequent purification steps or for protein storage in a stable storage buffer. Following the company's instruction, the dialysis membrane was soaked in water for about 20 mins in the first step and then equilibrated for 10 mins with the exchange buffer and next the sample was pipetted into the membrane tubing. The exchange buffer was approx. 200 to 500 times the volume of the sample and was slowly stirred for dialysis carried out overnight at 8°C.

3.3.8. Sample concentration

After each purification step, to achieve the concentration required for crystallization, the Pdx protein solutions were concentrated in the Eppendorf concentrators with a suitable molecular weight cut-off. The protein samples were applied on the concentrator and spun down for about 10 min at 3500-4000 rpm each until the desired protein concentration was reached. In between the two centrifugation steps the sample was resuspended, mixed with the diluted sample, and applied fresh again for the next run.

3.3.9. Size-exclusion chromatography (SEC)

Separation matrices with different pore sizes can be used to separate macromolecules by their size and shape. In the separation mechanism, small and globular-shaped particles penetrate more to the stationary phase of the matrix pores and hence hindered more while passing through the matrix. Whereas the bigger, as well as elongated particles, penetrate less into the matrix, therefore travel faster, and eluted earlier as compared to the small-sized particles. The latest ÄKTA purifier or the ÄKTA prime system were utilized during my PhD studies for size exclusion chromatography.

An ÄKTA fast protein liquid chromatography (FPLC) purification system (ÄKTA Purifier P-901; GE Healthcare, UK) connected to a UV detector and a fraction collector operated at 16°C for purification. The device was cleaned by pump wash purification and system washing with 50 mL water and next the column was washed with 1.5 CV of water and then equilibrated with 2 CVs of SEC purification buffer. The protein

separation parameters were optimized, controlled by the accompanying UNICORN software.

The concentrated sample of volume 4.5 mL was injected into the column through the injection valve and 5mL injection loop. At a constant flow rate, the absorption of the eluate was monitored at 214 nm and 280 nm by an installed UV lamp. Elution fraction peaks giving high UV absorption values were collected, stored, and analyzed with SDS-PAGE. Water and all buffers were filtered and degassed before applying to the ÄKTA purification system. The columns Hi-Load 16/60 Superdex 75, Hi-Load 26/60 Superdex 200, and Superose 6 increase 10/300, obtained from GE Healthcare (GE Healthcare, Uppsala, Sweden), were utilized for the purification of Pdx proteins.

3.3.10. SDS-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE is a technique utilized mostly in determining the size of a protein sample as well as purity of the protein sample under denaturing conditions. The principal ingredient is a polyacrylamide-based discontinuous gel, polymerized by a radical polymerization mechanism. The matrix acylamide is polymerized by a reaction driven by polydentate ligand TEMED with the addition of the radical initiator ammonium peroxide sulfate (APS). The smaller pore size of the polymerized acrylamide gels is the main difference to the agarose gels and can separate proteins up to 100 kDa. Further, it can also separate small-sized proteins up to 10 kDa, less than this mass are difficult to resolve because of the low binding ability of SDS [123]. For protein sample preparation SDS addition ensures the denaturation of the proteins into monomeric forms. SDS is an anionic detergent maintaining a net negative charge on the surface of protein within a wide pH range. The heat at 96°C, during sample preparation, breaks the intermolecular bonding present at the secondary and tertiary structure level of proteins. Further, the reducing agent like DTT destroys all the disulfide bridges present between the secondary structural elements. So, the uniform unfolded monomeric states of the protein molecules are accumulated by SDS ions, with the sulfate groups of the SDS directed outwards into the solution, and their negative charge overlay the intrinsic charges of the proteins. The negative charges on SDS further aid in destroying the complex structure of the proteins and therefore are strongly attracted toward the anode in an electric field. Because of the smaller pore size, polyacrylamide gels restrain larger molecules, whereas the smaller molecules migrate faster within the gel. Because the charge-to-mass ratio is

nearly the same among SDS-denatured polypeptides/proteins, the final separation of proteins is dependent almost entirely on their differences in relative molecular mass values.

In general, the gel consists of a narrow pore size separating gel consisting of 12% of acrylamide/bisacrylamide, as well as a focusing stacking gel with 4% of acrylamide/bisacrylamide. The ionic strength, as well as pH values of both gel layers, are different. The stacking gel is acidic (pH = 6.8) and when the proteins enter the basic separating gel (pH 8.8), the charge of glycine present in gel composition increases. Thus, the proteins are no longer focused between the two anions, hence separated by their size [124]. A list of all buffers utilized in gel casting and electrophoresis are summarized in table 7 and 8.

For SDS-PAGE evaluation of the purified Pdx proteins, samples were mixed with 2x SDS-sample buffer before denaturing them at 96°C for approx. 7-10 min. The heated and digested samples were spun down briefly, and 5 µL to 10 µL were loaded onto the gel. A protein marker of range 14-116 kDa (code: 26610, Thermo Fisher Scientific Inc., USA) was also applied to the gel to estimate the molecular weights of the Pdx proteins. The gels were run in a vertical gel electrophoresis chamber (Hoefer Inc, USA) and connected to an EV-231 power supply (Peqlab, Germany) and 1x electrode buffer was used for running the sample in the gel. A constant voltage of 180 volts for separation in stacking gel and 120 volts for separating gel and a current of 25 mA per gel was applied for approx. 1 h. Afterward, the gel was allowed to stain in a coomassie staining solution with continuous shaking for about 2 hrs to overnight. The next day the gel was de-stained by exchanging with the coomassie de-staining solution (20 % (v/v) acetic acid) several times until the protein bands could be distinguished clearly in the destaining solution from the background and documented by a color scanner (Epson).

3.3.11. Native gel electrophoresis

3.3.11.1. Clear/blue native PAGE

Native PAGE is used to analyze proteins in their physiologically folded state, where proteins are not denatured and unfolded. Here, the separation is based on their mass and charge allowing proteins to be detected in their oligomeric states.

Pre-cast polyacrylamide gels of 3-12% gradient for blue native PAGE were obtained from SERVA (Heidelberg) and used according to the instructions in the manual with

some modifications applying a mighty small-II PAGE chamber (Hoefer, USA). Furthermore, a standard protein native marker to approximate the molecular weight (SERVA electrophoresis, GmbH, Germany), sample buffer, as well as anode and cathode buffers, were purchased from SERVA (Heidelberg).

For Pdx proteins native-PAGE evaluation, protein samples were mixed with 2x native PAGE sample buffer (without SDS) in a 1:1 ratio and then loaded onto the gel, and electrophoresis was performed for approx. 2 h. The native-PAGE electrophoresis was carried out on the ice to avoid possible denaturation of the native proteins due to heat generated during the electrophoresis. The electrophoresis was started with a constant voltage of 50 V for 10 min until the proteins had migrated through the stacking gel. For migration of proteins through the separating gel, a voltage of 200 V was set and let run for 2 h. After electrophoresis, the gel was incubated and stained with coomassie brilliant blue R-250 for 2 h. Then it was transferred to the destaining solution (20% acetic acid) for 4 h, washed three times with water, and documented by a color scanner (Epson).

3.4. Biophysical Methods

3.4.1. Circular dichroism (CD) spectroscopy measurements

Circular dichroism (CD) spectroscopy is an excellent tool for evaluation of the secondary structure and folding of proteins, their interactions, and the complex formation, and to a lower extent information about their tertiary structures. By definition, circular dichroism is defined as the unequal absorption of left-handed and right-handed circularly polarized light [125]. CD spectroscopy is a very effective and sensitive technique characterizing small chiral organic molecules and large macromolecules like proteins. Primarily, the technique is employed in determining the secondary structure contents of proteins in the solution. It is also used to investigate the stability of proteins in the solution by measuring the melting point temperature after observing changes in folding properties.

To evaluate the secondary structure content and to score the overall folding of Pdx1, Pdx2, and the Pdx complex, CD spectroscopy investigations were performed utilizing a J-815 CD spectrometer (Jasco, UK) instrument equipped with a Peltier element. The CD spectra of the samples were recorded at 20°C and measured in 1 mm quartz cuvettes with a total volume of 100 μL and of 3-5 μM Pdx protein solutions with 10

measurements per sample. Pdx protein samples were centrifuged at 4°C for 20 min at 15,000 rpm, prior to the measurement.

CD spectra were measured for each Pdx protein sample over a wavelength range of 190-250 nm (far-UV spectrum) carried out with an interval of 0.1 nm, which were accumulated and averaged. The baseline was recorded for the corresponding sample buffer by the same procedure and subtracted from the protein signal to obtain the sample absorbance/ellipticity (θ). The ellipticity (θ) is defined as the difference in absorbance of clockwise and counter-clockwise circular polarized light as presented in **Eqn-2**.

$$\theta = \frac{180.ln10}{4\pi} \cdot (A_L - A_R) \qquad (2)$$

A_L and A_R are left circularly polarized (LCP) light and right circularly polarized (RCP) light, respectively. As the change in polarization is very small, so conversion factor (180/π) converting radians to degrees is used, θ is ellipticity measured in mdeg.

The results displayed by the Spectra Manager™ software provided by Jasco with the instrument as ellipticity [mdeg] against the wavelength, were converted to mean residue ellipticity (MRE) [deg cm^2 $dmol^{-1}$)] by **Eqn-3**, according to [126].

$$MRE = \frac{MRw \cdot \theta}{10 \cdot cd} \qquad (3)$$

Where θ is the measured ellipticity [mdeg], MR is the molar mass in $gmoL^{-1}$, *c* is the protein concentration [$mgmL^{-1}$], *d* is the path length through the sample measured in cm and W is the amino acid residues weight in a monomer.

The mean residue molar ellipticity was obtained by multiplying MRE with 1000 and plotted against the wavelength and the secondary structure contents were determined applying the Spectra Manager™ software [127]. The typical curves representing α-helical folding show minima at 208 and 220 nm and a maximum at 192 nm. The proteins with mainly β-sheet folding show a minimum at 215 nm and a maximum at 195 nm. Whereas random coil structures lead to a minimum between 190 and 200 nm and a maximum at 220 nm.

3.4.2. Dynamic Light Scattering (DLS) studies

Dynamic Light Scattering (DLS) is a very common biophysical and bioanalytical method applied in vast areas of life and materials science. DLS (also known as photon correlation spectroscopy or quasi-elastic light scattering) is used in determining the size distribution of small particles or macromolecules in solution. Further, it is used in investigating the reaction kinetics of macromolecules involved in a complex reaction [96]. In its basic principle, a monochromatic laser light hits the particles of the solution, is then scattered in all directions. The scattered light also interferes with the scattering of other particles while reaching the detector. Because of the Brownian motion, the position of the scattering varies all the time, and the scattering intensity fluctuates. These fluctuations are recorded and quantified over time by a photon-detector positioned at a 90° angle. The detector obtains information about the fluctuation in the light scattered by the particles, and the diffusion coefficient can be derived from the diffusion rate of the particles. The diffusion constant (D) is then used in determining the hydrodynamic radius (R_H) of the particles in solution in accordance with the Stokes-Einstein equation **(Eqn-4)** [128, 129].

$$R_H = \frac{k_B T}{6\pi \eta D} \quad (4)$$

R_H is the hydrodynamic radius that is determined by its dependence on the viscosity (η) of the solution and the diffusion coefficient (D). K_B is the Boltzmann's constant, and T is the temperature of the experiment.

To investigate the dispersity, size, and hydrodynamic radius information of the Pdx proteins, sample solutions were measured in a quartz cuvette using the Spectroscatter-301 (Xtal Concept, Hamburg, Germany) device applying a red-light class 3b lasers (λ = 660 nm; laser power 100 mW). The device has a photomultiplier tube detector. The device can visualize molecules in size range from 1 nm to approx. 6 µm, with a concentration range from 0.5 to approx. 100 $mgmL^{-1}$. For sample preparation, the Pdx protein samples with a concentration of 1, 2, 5, 10, 15, 20 $mgmL^{-1}$ were centrifuged at 8°C for about 120-180 min at 17,000 g to sediment the high molecular weight aggregates and analyzed the clear supernatant in the quartz cuvette at room temperature. For standard sample analysis 30 measurements for 30 seconds each, with no pause time in between, were performed. Measurement values were stored in a standard data (.dat) format for later analysis with the help of the provided SpectroSize 300 software.

3.4.3. Small-angle X-ray scattering (SAXS)

Small-angle X-ray scattering (SAXS) is a powerful biophysical technique involving X-rays used for the determination of the overall 3D shape and structural characterization of biological molecules in solution. It is applied for the characterization of both ordered and disordered proteins. It is employed for the size and low-resolution shape determination of molecules randomly oriented ranging in size from a few kDa to GDa. This technique also offers effective means to reaction kinetics of reactions involved in the seconds to minutes range [112].

For SAXS measurements, Pdx protein samples were prepared following the standard procedure of beamline provided [130]. The proteins were dialyzed overnight against a buffer, in which proteins were stable and had less background scattering. The dialysis buffer was kept for the SAXS buffer subtraction measurements as a blank. Three different concentrations for *P.vivax* and *S.aurues* Pdx1 i.e. 2 $mgmL^{-1}$, 5 $mgmL^{-1}$ and 10 $mgmL^{-1}$ and considering concertation dependent oligomerization of *P.vivax* Pdx2, two concentrations i.e. 2 $mgmL^{-1}$, 5 $mgmL^{-1}$ and *S.aureus* Pdx2 three different concentrations 2 $mgmL^{-1}$, 5 $mgmL^{-1}$ and 10 $mgmL^{-1}$ were prepared and measured. All SAXS measurements were performed at the EMBL beamline P12, PETRA III, DESY, Hamburg, Germany. The samples (20-100 μL) were exposed to the beam at 20°C for 0.50 seconds at a wavelength of 1.24 Å (with stepwise increasing protein concentration) at a detector distance of 3.1 m in separate experiments.

For complex formation, pure Pdx1 and Pdx2 proteins were mixed in a 1:1 ratio and incubated overnight. Prior to the SAXS data collection, the complex was purified by a Superose-6-increase column (10/300, GE Healthcare). Likewise, Pdx1, Pdx complex was also investigated at concentrations of 2, 5, and 10 $mgmL^{-1}$ applying an automated robotic sample changer and a Dectris 2D photon-counting detector (PILATUS-6M) placed at 3.1 m distance. The sample was exposed to the X-rays with a flux of 10^{13} ph/s and momentum transfer S($Å^{-1}$) ($S=4\pi\sin\theta/\lambda$, where 2θ is the scattering angle).

3.4.3.1. SAXS data processing steps

Scattering intensity data of the three Pdx proteins from *P.vivax* and *S.aureus* were collected, integrated, and averaged by applying the ATSAS online software suite (EMBL, Hamburg, Germany) separately [131]. The scattering intensity of the sample is proportional to the concentration of the sample. Higher concentration gives a better

signal-to-noise ratio of the sample data when subtracted from the respective buffer, providing more clear information about the sample size and shape. However, at higher concentrations, the distances between the individual particles decreases, aiding interparticle interference and thus contributing to the scattering intensity pattern. Therefore, samples were applied in the concentration range, and data were scaled and averaged in the final steps of data processing.

For initial data processing, PRIMUS was used for Guinier analysis [132]. First, the individual scattering profiles of Pdx proteins at one (lowest) concentration were collected and inspected independently for interparticle interference, aggregation, or radiation damage. Then, the scattering curves were scaled and averaged to one another after showing no difference between the profiles, and the Guinier analysis plots were calculated. These calculated Guinier plots can be used to determine the radius of gyration (Rg) by knowing that linear Guinier plots indicate the monodispersity of the sample. The pair distribution functions *P(r)* indicating maximum diameter (*Dmax*) of proteins, and forward scattering *I*(0) were computed in GNOM [133] and PRIMUS [132]. The obtained data were compared with the values provided by an online SAXS data processing server named SAXSMoW [134].

Further, the obtained R_g values of the proteins were compared with the R_H value determined by DLS respectively with the help of the following equation **(Eqn-5)** [135]. The comparison represents a suitable approximation about the shape information of particles in solution.

$$\rho = \frac{R_g}{R_H} \qquad (5)$$

ρ is the shape factor
R_g is the radius of gyration obtained from SAXS
R_H is the hydrodynamic radius obtained from DLS

Kratky plots ($I_{(S)}S^2$ vs S) and Porod-Debye plots ($I_{(S)}S^4$ versus S) provide preliminary shape and flexibility information, were generated as described [136] and molecular masses (MM) information were obtained from SAXSMoW [134]. The low-resolution *ab-initio* models were generated from the scattering curves by dummy bead modeling programs DAMMIN and GASBOR considering symmetry option P1 for Pdx2 and P6 for Pdx1, and the complex. Whereas the *S.aureus* Pdx complex was investigated for 3D

modeling applying different symmetries options [137, 138]. Dummy bead modeling programs utilize simulated annealing to construct a densely packed beads model based on experimental scattering intensity data. Five DAMMIN and GASBOR runs were performed for each of the proteins and the results were well superimposable with each other. Rigid body models were constructed from scattering data and 3D-models were calculated by the program SASREF [139], considering the same symmetry options as those used for the dummy models.

ATSAS program CRYSOL allows comparison as well as the superposition of the obtained scattering data for proteins to the known high-resolution homologous PDB structures. Therefore, theoretical scattering curves of homologous structures deposited at the protein data bank were calculated, compared, and fitted to the experimental scattering data of Pdx proteins utilizing CRYSOL [140].

3.4.3.2. Size exclusion chromatography (SEC)-SAXS

Biological SAXS has become a very streamlined technique for structural parameters characterization involving conformational changes of proteins and macromolecular complexes in solution [141]. However, samples likely aggregating and oligomerization, partially associating and dissociating complexes, 3D solution structural interpretation becomes difficult as SAXS structural analysis requires monodisperse sample solutions [142]. However, this enigma was overcome by a complementary SAXS technique called size-exclusion chromatography-SAXS (SEC-SAXS) [143]. It is effectively used to separate the polydisperse solutions based on the molecular weight of the samples. It continuously measures the sample while it elutes from the SEC column and exposing the eluted samples to the X-ray beam.

Since, in some proteins oligomerization, while in others complex formation is concentration-dependent. Therefore, information about the accurate concentration determination while eluting from the SEC column is important. For this purpose, SAXS is combined with ultraviolet-visible light (UV-vis) and refractive index detectors for quantifying the concentration of the eluting sample. Further, the set-up also combines with right-angle light scattering (RALS). The RALS detector provides the estimation of the molecular weight of the eluted samples. The MW estimate integrated over a peak volume, allows one to independently assess the oligomeric states of the eluting fractions. The combination of SAXS with other biophysical and bioanalytical

techniques provides full automation of the SEC-SAXS data processing and analysis of the sample mixture [144].

3.4.3.3. SEC-SAXS of *P.vivax* Pdx2 and *S.aureus* Pdx proteins

Considering the time and concentration-dependent oligomerization tendency of *P.vivax* Pdx2 (discussed in detail in the Results section of the thesis), online size-exclusion chromatography was applied prior to the SAXS data collection by applying 7 $mgmL^{-1}$ of *P.vivax* Pdx2. For SEC run, Superdex-75-increase column (5/150, GE Healthcare), which was pre-equilibrated with a buffer of 20 mM Tris-HCl (pH 8.0), 200 mM NaCl, 1 mM EDTA, 2 mM DTT. The buffer also includes 3% glycerol to aid solution stability and reduce the risk of radiation damage. The obtained eluted sample solution was directly subjected to X-ray scattering data collection. For sample exposure, the flow cell method was used, where the sample flows continuously and is exposed to the beam for the duration of the experiments.

Observing the different oligomerization behavior of *S.aureus* Pdx1 revealing hexamer and other low molecular weight multimeric species in mass photometry investigations (details are provided in the Results section of the thesis). SEC-SAXS was employed to get structural insights into the sample eluted peaks from the SEC column. Subsequently, *Sa*Pdx2 and Pdx complex were also investigated by SEC-SAXS to observe and confirm their multimeric states. The Pdx proteins were concentrated 7 to 10 $mgmL^{-1}$, filtered and centrifuged before loading onto the respective SEC column. A concentrated sample of volume 50-80 µL was loaded to the Superdex-75-increase column (5/150, GE Healthcare) after column calibration and SEC runs were performed at a flow rate of 0.5 mL/min at room temperature.

3.4.4. Nanoparticle tracking analysis

Nanoparticle Tracking Analysis (NTA) is one of the few methods to visualize and also measure the nanoparticles suspension in the size range from a few nm to 1000 nm based on tracking of Brownian motion [145]. Brownian motion or Brownian movement represents the random movement of solute particles in the suspension. It is the fluctuations in a particle's position inside a solution at one position which is followed by relocation to another position with the same or similar fluctuations. For detection and tracking, the solute particles must scatter light with an intensity level above a certain threshold value. The amount of light scattered by the solute particle is proportional to

the scattering cross-section of the particle exposed at one time to the incident light and the intensity of the incident light source [146].

NTA track and determine the Brownian motion of the particles in the homogenous, as well as in the heterogeneous sample by visualizing and illuminating the individual particles tracked by the help of a laser beam. The scattered laser light is recorded with a light-sensitive camera, arranged at a 90° angle to the irradiation plane. This arrangement allows the detection and tracking of the Brownian motion of particles. Further, the size of each individually tracked particle is calculated using the Stokes-Einstein relation [145], solving the particular particle diameter (d), shown in **Eqn-6**.

$$D_t = \frac{k_B \cdot T}{6 \cdot \pi \cdot \eta \cdot R_H} \qquad (6)$$

Where Dt is the diffusion coefficient of the particle in the solution, K_B is the Boltzmann's constant, T is absolute temperature, η is the viscosity of the solution without the solute and R_H is the hydrodynamic radius of the solute particle to be investigated.

Since *P.vivax* Pdx2 exhibited reversible oligomerization tendency with the size range of the oligomeric particles from 50-200 nm. To observe this behavior, the oligomeric Pdx2 was allowed to mix with the dodecameric Pdx1 and subsequently observed for diameter shift. Pdx1 (0.25 ml) was added to the oligomeric Pdx2 (0.25 ml) inside the sample chamber. In the NTA experiment, the NanoSight LM10 (NanoSight Ltd., United Kingdom) system with a laser beam of wavelength 405 nm for sample observation was utilized. Particles were visualized using a conventional optical microscope, connected to a video camera for tracking the movement of the particles in solution. The sample was measured for 60 seconds at different reaction times at a rate of 25 frames per second with manual shutter and gain adjustments. The "single shutter and gain mode" was used to capture the Pdx proteins and analyzed for size and diameter by the NTA software version 3.2 (NanoSight Ltd).

3.4.5. Mass photometry studies

Sample purity is the prerequisite for any successful structural investigating technique like macromolecular X-ray crystallography and cryo-electron microscopy. Mass photometry is the technique accurately characterizing the sample homo or heterogeneity at high resolution with minimum material cost and in a matter of minutes [147]. Mass photometry (MP) works on the principles of interference reflection microscopy [148].

The minute quantity of light scattered by each single sample molecule present in the solution can be detected much reliably and correlated directly to the molecular mass of the sample [147].

Since, protein characterization methods such as SDS-PAGE, size-exclusion chromatography (SEC), and dynamic light scattering (DLS) characterize the sample composition [149], and each has significant limitations. SDS-PAGE reveals molecular weight information with stoichiometry or interactions with sample concentration limitations. SEC works on the basis of Stokes radii, where the solvent largely affects the sample size based on the mobility/density of the solution and hence does not report the actual molecular weight. DLS has limited mass accuracy and also resolution with concentration limitation [147]. Whereas, MP measures the mass of solute in solution after quantifying the light scattered from the individual biomolecules in solution with around 5 kDa deviation between the sequence and the measured mass [150]. The method also allows the direct detection of proteins when they assemble into functional complexes, oligomeric or misfolded aggregates [151].

Pdx1 protein from other plasmodial species is reported for concentration-dependent oligomerization and subsequent aggregation. Whereas bacterial Pdx1 is reported in hexamer-dodecamer equilibrium (elaborated in detail in the Introduction and Discussion section of the thesis). To unveil this enigma of concentration-dependent and species-specific oligomerization, Pdx1 from *P.vivax* and *S.aureus* is investigated by mass photometry at different concentrations, ranges from 10 nM to 1 µM. This concentration of Pdx1 protein contains 10^9 to 10^{11} Avogadro's number of protein molecules, which is in the optimal measuring range for the MP instrument.

In MP investigation simple microscope coverslips are used, cleaned by sonication in 50% isopropanol followed by clean air stream drying. Protein and buffer samples were centrifuged and filtered before measurement. Immediately, before the mass photometry measurements of Pdx proteins, protein stocks were diluted directly from the buffer stock solutions to the final measuring concentration of 10 nM to 1 µM and then measured for 30-60 seconds. Measurements were performed with the Refeyn One^{MP} mass photometer (Refeyn Ltd. UK) instrument with a light wavelength of 525 nm and molecular mass detection range from 40 kDa to 5 MDa. Afterward, the recorded images were processed and analyzed using Discover^{MP} (Refeyn Ltd. UK) software.

3.4.6. Negative staining Electron Microscopy (EM) investigations

Negative stain electron microscopy is an established technique, used for contrasting a thin sample specimen with an optically opaque fluid [152]. In this technique, the background is stained by a staining reagent and leaves the sample specimen undisturbed and being visible. Whereas, single-particle electron microscopy (EM) technique provides structural information of a large variety of biological molecules in their native states, ranging from proteins to large functional macromolecular assemblies. It worth mentioning the year 2008, which is a landmark year for single-particle electron microscopy. This year the density maps have been produced for the first time at a resolution able in tracing proteins backbone and building 3D atomic models, revealing the internal structures [153].

In biological assemblies investigation, macromolecule particles are adsorbed on an electron transparent sample support grid (called EM grid) in the form of a thin amorphous film. The thin film of the sample distributed over the entire grid almost evenly is then investigated by a transmission electron microscope to reveal its external structural details [153, 154]. Since electrons interact with the matter very strongly. Therefore, the electron path from emission to the sample and camera in the microscope must be kept under high vacuum conditions to avoid any unwanted scattering by the air or gas molecules in the chamber. Accordingly, the considered EM specimen must be in the static position for imaging. Further necessary specimen preparation techniques are required for dehydration or stabilized hydrated biological samples in a vacuum [155].

In my PhD studies, the negative stain EM investigations were utilized for the qualitative examination of the purified Pdx proteins prepared and to depict the dynamic behavior of the Pdx complex with oligomeric subunits Pdx1 and Pdx2. The EM staining procedure was performed by first preparing the negative stain EM grids that will support the sample. Specimens were prepared for EM investigations following a conventional negative staining procedure [156]. The carbon-coated grids were discharged and made the grid hydrophilic by the Quorum GloQube glow discharger (East Sussex, UK) for 30 seconds using a current of 25 mA. A filter paper was placed in a fuming hood, and a glow discharged grid was placed on the paper.

The sample solutions with concentrations between 25-30 µg mL^{-1} (molar concentration of 750-900 nM), of Pdx1, Pdx1 in complex with Pdx2 monomers, and Pdx1 in complex

with Pdx2 oligomers were used in EM studies. The three samples were prepared in 20 mM Tris pH 8.0 buffer containing 200 mM NaCl and 5 mM MTG (monothioglycerol) and applied on separate grids. 4 µL of samples containing around 10^{10} to 10^{11} Pdx protein molecules, were applied on the glow-discharged carbon-coated side of the copper grid and incubated at room temperature for 30 seconds for maximum absorption of protein on the grid. The remaining undissolved sample was discarded from the grid. The waiting time and the time length of glow discharging of grids, affect the protein concentration adsorbed on the grid. Very diluted samples require longer adsorption times and concentrated ones require less. Next, the grid with absorbed protein molecules was washed two times with water and blotting off in between to remove any unbound sample stick and hanging the grid. Subsequently, the grid was dipped in the 2% uranyl-acetate solution for negative staining in four successive steps. After letting it dried, grids were ready and saved for EM investigations.

In the next step, EM grids were loaded individually in the single tilt standard specimen holder of Talos L120C (Thermo Fisher Scientific-FEI) electron microscope at 120 kV accelerating voltage and a LaB_6 electron source filament with a CETA camera. The specimen was focused, and a desired defocus with maximum contrast was set. Focusing was achieved either manually using the contrast method or by using the CCD camera to capture live images and compute live Fourier spectra. Afterward, in the final step, images were taken using a magnification of 71,000× to 92,000×.

3.5. Crystallization, diffraction data collection and model building

3.5.1. High-throughput screening for crystallization conditions

Before crystallization, fresh purified Pdx protein solutions were concentrated and centrifuged for 2 hrs at maximum speed. Subsequently, the sample monodispersity was confirmed by the DLS investigation. Before proceeding to crystallization, protein concentration considered to be optimal for crystallization was investigated by pre-crystallization test (PCT) (Hampton Research) with the protocol of the company's instructions followed.

For initial crystallization trials, vapor-diffusion experiments were set up in either 2-well MRC 96-well sitting drop crystallization plates (Molecular Dimensions, UK) or MRC maxi 48-well crystallization plates (Jena Bioscience Germany). For screening of various buffer conditions, different high-throughput screens were utilized (shown in Table-10)

using the Pipetting robot Honeybee 961 (Genomic Solutions, UK). In high-throughput screening 300 nL of the Pdx protein solution was mixed with 300 nL of the precipitant solution and equilibrated against 50 µL of reservoir solution in sitting drop vapor diffusion condition using MRC-2 96-well sitting drop crystallization plates (Molecular Dimensions, UK). Plates were incubated at a specific constant temperature of 20°C and 4°C in a RUMED 3001 incubator (Rubarth Apparate GmbH, Germany) in individual experiments.

3.5.1.1. Protein crystal visualization

After a few successful crystallization hits obtained from high-throughput screening experiments. The protein and salt crystals differentiation were performed by analyzing the protein drops with the Crystal Score device (Diversified Scientific Inc, USA). The device detects a specific UV fluorescence of protein crystals and differentiates them from salt crystals.

Subsequently, drops containing crystals were also mixed and incubated with Izit Crystal Dye (Hampton Research, USA). It is a small molecule dye that is small enough to penetrate the solvent channels present in the protein crystals. In its basic working principle, it binds to the negatively charged chains of amino acid molecules utilizing the intermolecular spaces arranged in the ASU and unit cells of crystal, thereby coloring the protein crystals. Whereas it is too big in its structure to penetrate salt crystals. The recommended 1:10 dilution of Izit Crystal Dye in ddH_2O was prepared, and an amount of 0.5 µL was pipetted directly onto crystals. The drop was then incubated for 30 minutes to 1 h and afterward inspected under a light microscope.

3.5.1.2. Crystallization conditions optimization

After obtaining initial crystallization conditions and protein crystals confirmation, the crystallization conditions were optimized by varying the precipitant and to less extent, the protein concentration around the condition found in the initial screening. The promising crystallization conditions were refined by using the sitting drop vapor diffusion approach and crystallization under oil was performed as well in 96-well Terazaki plates (Nunc, Denmark) applying the Oryx 4 robot (Douglas, UK). For crystallization under oil, paraffin oil (Applichem, Germany) was added to the Terazaki plate (Nunc, Denmark) to a level that fills all the wells. Next, 1 µL of precipitant, followed by protein solution were pipetted into each well, while working using a

microscope. The oil covered the droplet and allowing controlled diffusion of the solvent molecules.

Further, for crystallization condition manual optimization, the hanging drop vapor diffusion approach was applied utilizing 24-well Linbro plates (Jena Bioscience Germany). In the hanging drop method, 2 µL of protein solution and 2 µL of the precipitant were mixed in the droplet, and the reservoir position in the plate was filled with 500 µL of precipitant solution. In parallel, crystallization conditions were also optimized by the sitting drop vapor diffusion method using 48-well MRC maxi sitting drop plates (Jena Bioscience Germany). In this method, crystallization drops were made by adding 1 µL of protein solution plus 1 µL of a precipitant solution, and the reservoir was filled with 50 µL of the precipitant solution. In both experiments, plates were stored in 4°C and 20°C incubators.

The additive screening kit from Jena Bioscience (Germany) was also applied for crystallization of the Pdx complex, where MRC maxi 48 well plates were used with 500 nL of protein solution mixed with 500 nL of precipitant solution. Further, 200 nL of the respective additive solution was added with 60 µL of reservoir solution in each well. The plates were sealed and stored at 4°C and 20°C.

After successful optimization of crystallization conditions, the final optimized condition for *P.vivax* Pdx1 was at a protein concentration of 20 $mgmL^{-1}$, and precipitant 0.1 M Tris, 200 mM $MgCl_2$, 19 % ethanol. Pdx1 from *S.aureus* crystallized by sitting drop crystallization method at a concentration of 10 $mgmL^{-1}$ using the Morpheus screen with condition details (0.09 M NPS (sodium nitrate, sodium phosphate dibasic, ammonium sulfate) 0.1 M Tris (base); BICINE P^H 8.5, 30 % v/v PEG 500 (MME) and PEG 20000).

3.5.2. Diffraction data collection

Diffraction data of Pdx proteins were collected at beamlines P11 PETRA III, DESY (Hamburg, Germany) and EMBL beamline P13 PETRA III, DESY (Hamburg, Germany). Both beamlines are using Pilatus 6M-F (25 Hz) detectors (Dectris Ltd., Baden, Switzerland) [115].

Single crystals were cryoprotected by transferring them into a reservoir solution with additional 20% glycerol for *P.vivax* Pdx1 and 30% glycerol for *S.aureus* Pdx1.

Subsequently, cryo-cooled in liquid nitrogen and stored in pucks (Uni-pucks for P11 and SPINE-pucks for P13) for safe transfer to the beamlines.

Diffraction data were collected for *P.vivax* Pdx1 and *S.aureus* Pdx1. Unfortunately, most of the *P.vivax* Pdx complex crystals did not diffract, whereas only a few crystals show diffraction at a very low resolution of around 10 Å. All standard data collections were carried out at 100 K and 12 keV, with flux intensity of approximately 2×10^{13} photons sec^{-1}. The exposure time was set to 40 ms with a beam of size 100×100 μm^2 or 200×200 μm^2 was used.

3.5.3. Data indexing and data reduction

The Pdx1 data were processed and indexed with the automated software XDSAPP [120]. The individual diffraction patterns were inspected and cut at a resolution above where the overall R_{merge} rose above 7.5%, or the I/σ value dropped below 1.2.

3.5.4. Structure solution and refinement

3.5.4.1. Molecular Replacement

The mtz file generated in the data reduction step was used further for molecular replacement, the first step towards 3D structure solving. In molecular replacement, the phase information is obtained utilizing the coordinates of a reported pdb model with appropriate homology to the structure in the investigation and to be solved. This also requires the FASTA sequence of the investigating structure to compare during the run. For calculation of the amplitudes (F calc) and the phases (Φ calc), solved model coordinates are positioned in the unit cell by both rotation and translation functions. Both the rotation and translation of molecules in the unit cell and ASU are solved by comparing them with the Patterson map of the known PDB structure.

With the help of the Patterson map, the intramolecular vectors (orientation of the arranged molecule) are used in the rotation function. The Patterson map of the structure to be solved is compared to the known structure Patterson map in different orientations until both the structure vectors have a similar orientation. Next, after obtaining the orientation of the molecule, the translation function, based on the intermolecular vectors, then places the molecule in the asymmetric unit at correct coordinates to get phase information. Then the phases derived from the model and the observed

experimental amplitudes (Fobs) are used to calculate the novel electron density map [118].

Pdx1 protein structures were solved by the Phenix 1.15.2-3472 program suite [157] and Phaser-MR was utilized for molecular replacement [158]. Phaser utilizes the maximum likelihood algorithms first for rotation and then for translation of the rotated model determination in the unit cell. The output of the Phaser run provides a PDB file containing the placed model, also an mtz file including coefficients for the electron density created by utilizing the placed model and the experimental F_{obs}. For solving *P.vivax* Pdx1 structure by molecular replacement, the reported homologous model from *Plasmodium berghei* (PDB code: 4ADT) [69] with 85% sequence identity was used as a search model to obtain the phases. For *S.aureus* Pdx1, the reported homolog Pdx1 from *Bacillus subtilis* (PDB entry 2NV1) with 80% identity [90] was used as a starting model applying Phaser-MR [158].

3.5.4.2. Structure refinement and validation

After initial macromolecular structure models are obtained from Phaser-MR, they were further processed by refinement using the Phenix-refine program [159]. Refinement consists of minimizing errors in the geometry of bond lengths and angles using different restraints, which remains after real space model building [117]. Based on these parameters a model is refined against the experimental data to obtain the best fit between the observed structure-factor amplitudes (reflection intensities) from the native data set (F_{obs}) and the calculated structure amplitudes from the model to be refined (F_{calc}). It is quantified by the R_{factor} values (R_{work}) indicating the magnitude of the difference between both the models **(Eqn-7)** [118].

$$R_{factor} = \frac{\sum [F_{obs} - F_{calc}]}{\sum F_{obs}} \quad (7)$$

If the observed and calculated structure-factor amplitudes agree with each other, the R factor is small and for a perfect agreement, R is zero (which is practically not possible). On the other hand, when both reflections intensities do not agree, R with a maximum value of one. In general, for proteins, the acceptable R values are 0.2 to 0.4 after the best placement of both model intensities. Further, for cross-validation, 5% of the

experimental data are utilized for calculating the R_{free} value, using the same equation as for R_{work} (equation-7) [118].

The low-resolution data collected for *S.aureus* Pdx1 revealed that the N terminal helix, C terminal flexible loop region, and helix α2′ were not convincingly visible after refinement. These regions were manually modeled in the electron density maps with the help of a program Coot [160]. Further, the quality of the refinement parameters and coordinates were verified by the online program REDO [161] and the online PDB validation server [162].

4.0. Results

4.1. Expression optimization of Pdx proteins

Pdx1 and *pdx2* genes of *P.vivax* and *S.aureus* with restriction sites (*NdeI* and *XhoI*) in expression vector pET31b(+) (BioCat GmbH) were transformed and expressed in different bacterial strains and their expression were optimized. After expression optimization, two bacterial strains BL21-CodonPlus-(DE3)-RIL and BL21(DE3) translate Pdx proteins optimally, the former with maximum over-expressed proteins. Therefore (DE3)-RIL was selected for Pdx proteins production for all future experiments. In the next step, the protein expression inducer IPTG (Isopropyl β-d-1-thiogalactopyranoside) concentration optimization was performed. The expression vector was induced at the IPTG concentration range from 0.1 to 1.0 mM, and 0.5 mM represents the optimal inducing IPTG concentration **(Figure 13a, b)**. Successfully optimized overexpressed C-terminal His_6-tagged Pdx1 and Pdx2 proteins were brought to SDS-PAGE, which confirms the molecular weights of 35 and 25 kDa respectively, indicating the presence of both the overexpressed proteins in the culture **(Figure 13a-c)**.

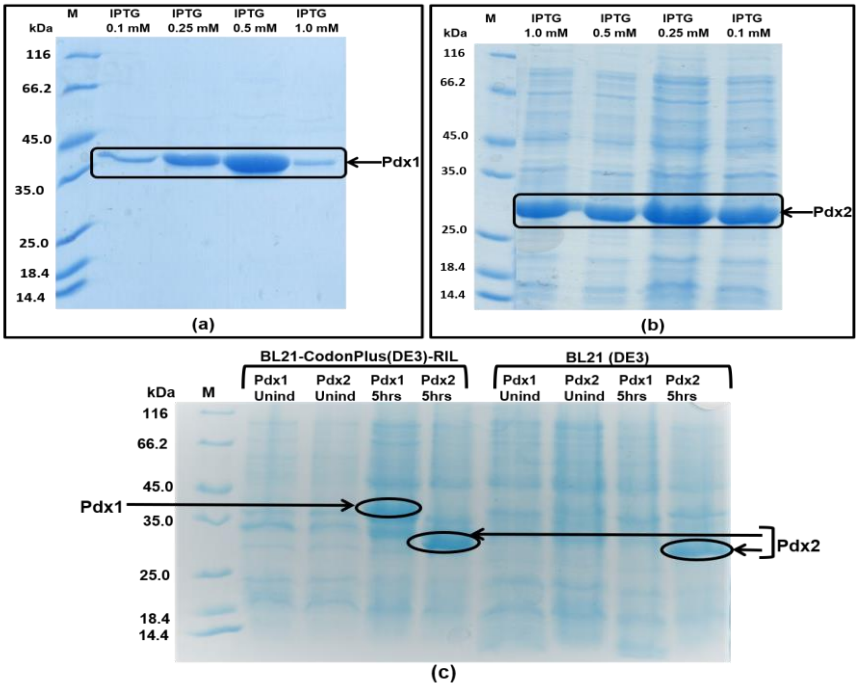


Figure-13. Expression optimization of *P.vivax* and *S.aureus* Pdx1 and Pdx2 proteins.
(a) and **(b)** are the IPTG concentration optimization in bacterial strain (DE3)-RIL, both Pdx1 and Pdx2 induces in a good amount at 0.5 mM IPTG concentration. SDS-PAGE indicates a broader band at around 35 kDa for monomeric Pdx1 and around 25 kDa for Pdx2 monomer **(c)** Bacterial strain selection for expressing a reasonable amount of Pdx proteins. In SDS-PAGE respective bands of Pdx proteins after comparing with the uninduced cultures indicate Pdx1 and Pdx2 expression. After 5 hours of IPTG induction, both proteins were expressed in a high amount, confirmed by their respective molecular weight band positions.

4.2. Pdx proteins purification

4.2.1. Pdx proteins Ni-NTA purification

After successful expression optimization, both Pdx proteins containing hexa-histidine tag were purified by Ni-NTA affinity chromatography as an initial purification step. *E.coli* cells with overexpressed proteins were lysed by sonication and clarified by centrifugation and subsequently applied on Ni column for purification. The flow-through was collected to confirm the proteins binding to resin, washing fraction was collected for imidazole concentration optimization with loosely bound Pdx proteins leaked out and followed by elution fractions collection with elution buffer containing 250 mM imidazole concentration. The eluted and collected fractions were evaluated by SDS-PAGE for Pdx proteins identification, as shown in **figure 14a, b** for *P.vivax* Pdx proteins, and **figure 14c, d** for *S.aureus* Pdx proteins.

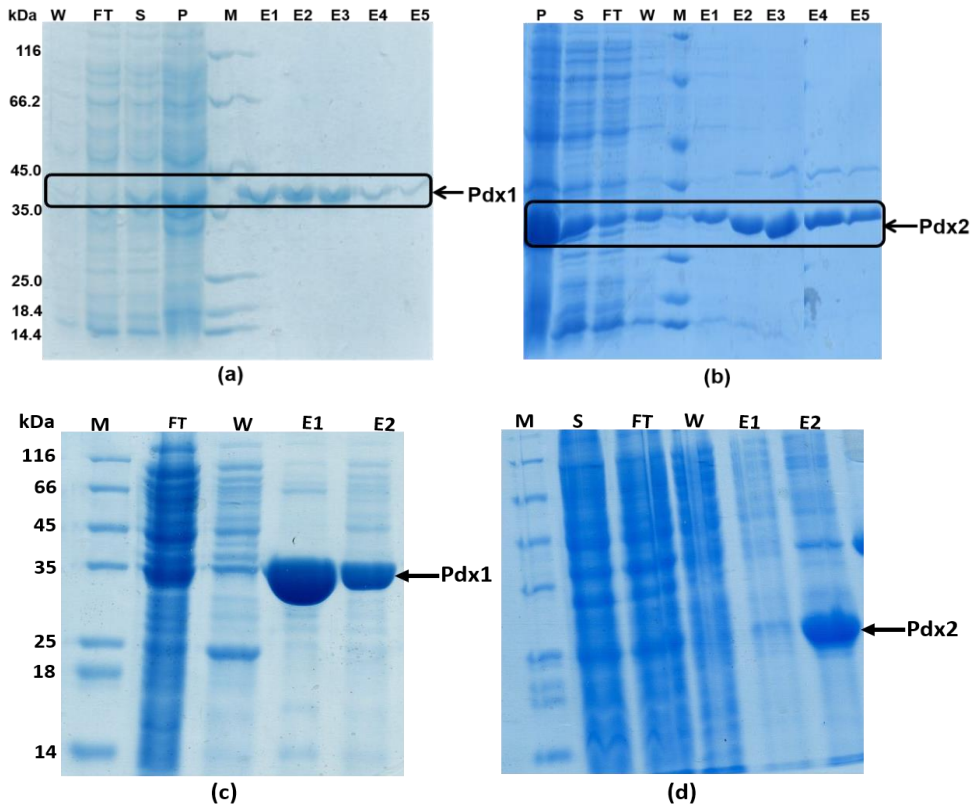


Figure-14. Ni-NTA purification optimization of Pdx1 and Pdx2 proteins.
P.vivax **(a)** Pdx1 **(b)** Pdx2 proteins were eluting gradually in the elution buffer, indicating a strong binding interaction with Ni-resin. As some of the *P.vivax* Pdx2 protein remains in pellets, so a bit harsh lytic conditions were applied, but it was leading to its aggregation. Therefore, mild extraction conditions were applied for *Pv*Pdx2 with compromise on yield. *S.aureus* **(c)** Pdx1 **(d)** Pdx2 proteins eluted in elution buffer very quickly as seen from the bands appear in SDS-PAGE at around 33 and 25 kDa, respectively.

4.2.2. *P.vivax* Pdx proteins SEC purification

The Ni-NTA purified Pdx proteins were applied on a size exclusion chromatography (SEC) column, individually. In SEC purification, *P.vivax* Pdx1 was eluted at a volume of approximately 160-170 mL by applying a Superdex S200 26/60 column for purification. The SEC elution profile for *Pv*Pdx1 is shown in **figure 15a**. The eluted fractions collected from the corresponding peaks 1 and 2 were evaluated by SDS-PAGE **(Figure 15b)** and further forwarded to native-PAGE evaluation **(Figure 17c)** showing that Pdx1 is eluting at approximately 400 kDa. This indicates that *P.vivax* Pdx1 is a dodecamer in solution with a calculated monomer molecular weight of 35 kDa [122].

In the SEC elution profile, *P.vivax* Pdx2 was eluted at approximately 230–256 mL **(Figure 15c)**, with its corresponding peaks 1 and 2 evaluated by PAGE (**Figure 15d** and **17c**), indicating a monomer and different multimers in solution, respectively. The monomeric form of Pdx2 was eluted with approx. 26 kDa following the calculated monomer molecular weight [122].

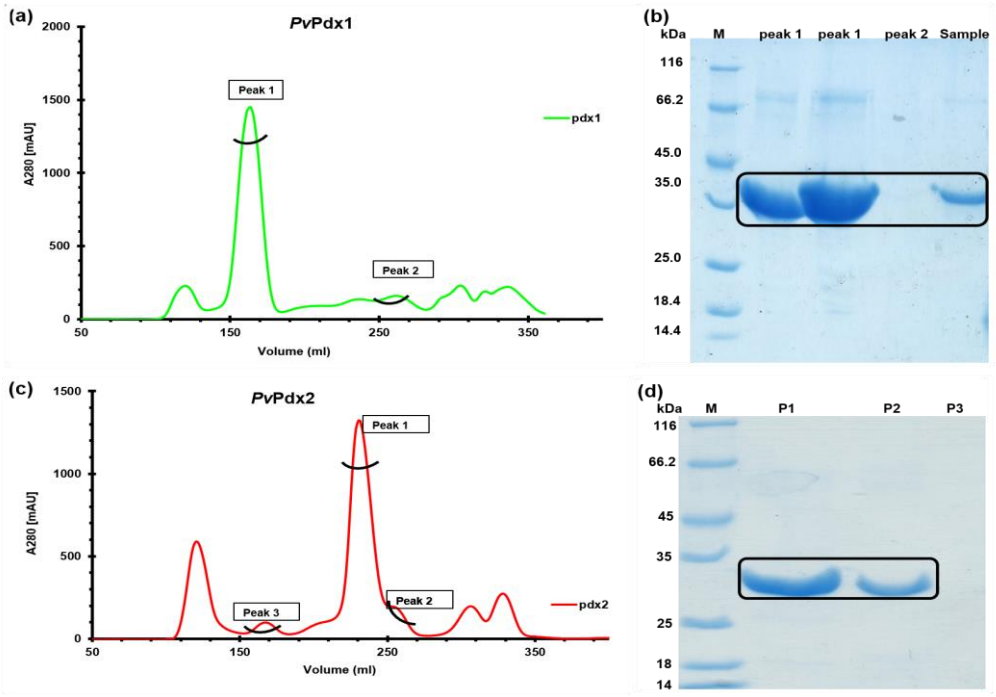


Figure-15. Size-exclusion chromatography (SEC) elution profile of *P.vivax* Pdx proteins.
(**a**) *P.vivax* Pdx1 and **(c)** Pdx2 purification using a Superdex 200 26/60 column and (**b**) SDS-PAGE evaluation of Pdx1 indicates Pdx1 in eluting peak1 and **(d)** Pdx2 elutes in peak 1 containing different oligomeric states whereas peak 2 indicate small amounts of eluting monomeric Pdx2 [96].

4.2.3. *S.aureus* Pdx proteins SEC purification

S.aureus Pdx proteins were purified using the SEC purification column Superdex S200 26/60 with the same purification conditions used for *P.vivax* Pdx proteins. *S.aureus* Pdx1 eluted at approx. 175-180 mL and its elution profile is shown in **figure 16a**. Whereas, Pdx2 was eluting with a prominent elution peak at around 250 mL **(Figure 16b)**. The eluted fractions were collected and evaluated by SDS-PAGE, indicating that the Pdx1 (P1) eluted at around 35 kDa and Pdx2 (P2) eluted at mass approx. 20-25 kDa, as shown in **figure 16c**.

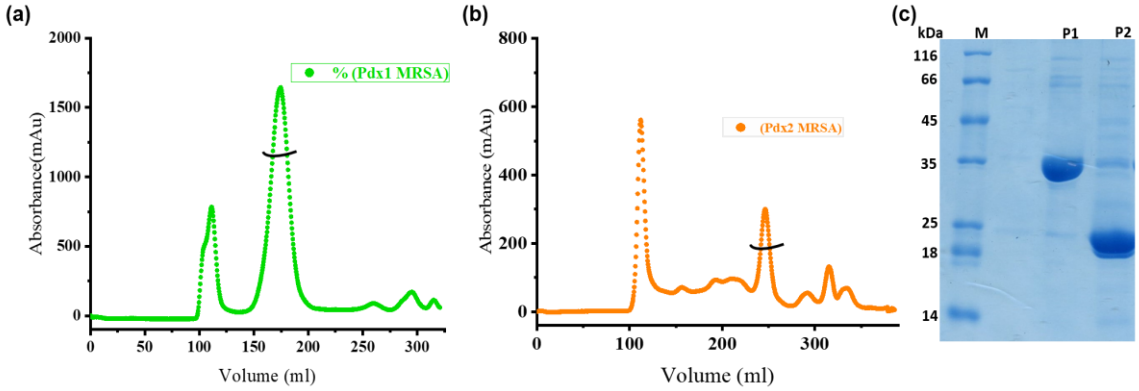


Figure-16. SEC elution profile and SDS-PAGE evaluation of *S.aureus* Pdx proteins.

(**a**) *Sa*Pdx1 SEC elution profile **(b)** *Sa*Pdx2 elution profile from Superdex 200 26/60 column (**c**) SDS-PAGE evaluation of *S.aureus* Pdx1 and Pdx2 as P1 at 35 kDa and P2 at 20-23 kDa bands, indicating their presence in their respective SEC eluting peaks.

4.2.4. Pdx complex purification

For Pdx complex formation and purification, the purified Pdx1 and Pdx2 proteins were mixed in a 1:1 molar ratio at a concentration of 0.5 mM in presence of 10 mM L-glutamine. The mixture was incubated overnight at 8°C and centrifuged at 10000 g for 1 hour and applied to a gel filtration column. *P.vivax* Pdx complex was purified through Superose 6 Increase 10/300 (GE Healthcare) column with its elution profile is shown in **figure 17a** and *S.aureus* Pdx complex is shown in **figure 17d**. *P.vivax* Pdx complex was eluted at approximately 15 mL volume range with an eluted peak corresponds to the Pdx1-Pdx2 complex **(Figure 17a)**. This chromatography fraction was assessed via SDS-PAGE, showing two bands at 35 and 26 kDa, corresponds to the reduced monomeric Pdx1 and Pdx2 (**Figure 17b** lane C, D). The purified Pdx complex was further characterized by native PAGE, with a band corresponding to a molecular weight of 650–700 kDa (**Figure 17c,** lane D) was observed. The *S.aureus* Pdx complex SEC

elution peaks, obtained by applying the aforementioned Superose column, were accessed by SDS-PAGE, where peak-1 and peak-3 were un-reacted oligomeric Pdx1 and Pdx2, respectively **(Figure 17d,e)**. Whereas peak-2 eluting at around 19 mL corresponds to the *Sa*Pdx1-Pdx2 complex, with a broad Pdx1 band and a narrower Pdx2 band in the SDS-PAGE **(Figure 17e)**.

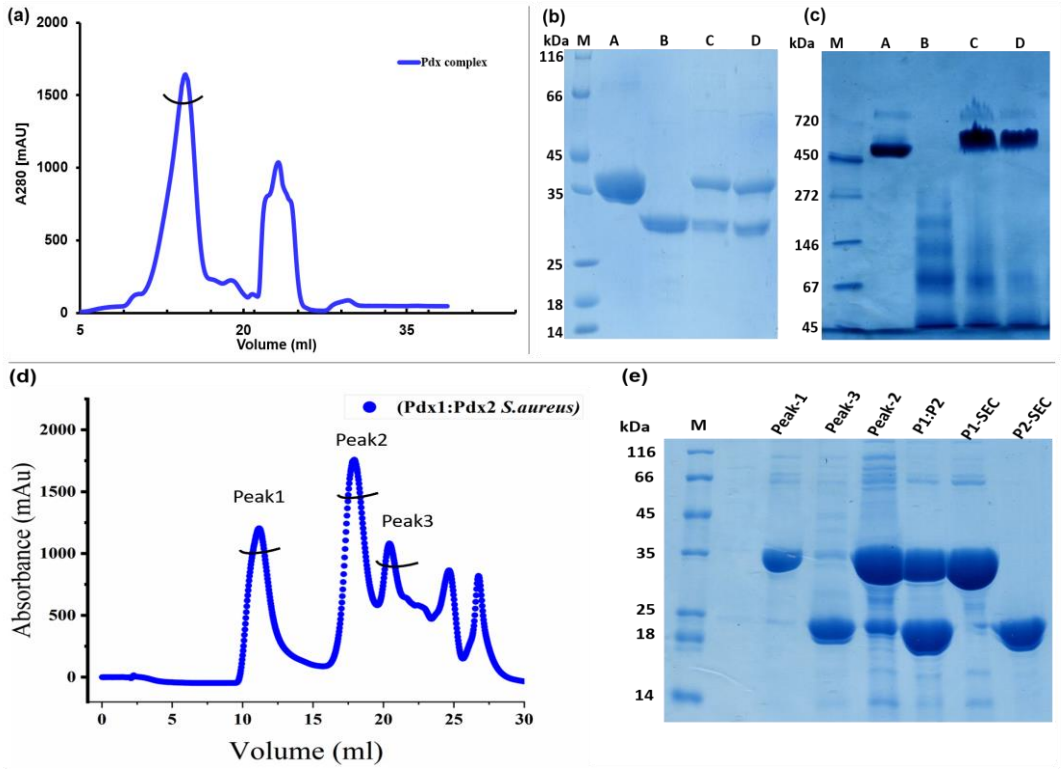


Figure-17. SEC elution profiles and PAGE analysis of Pdx complex.
(**a**) SEC elution profiles of *P.vivax* Pdx complex. **(b)** SDS-PAGE evaluation of *Pv*Pdx complex. Two bands of the corresponding proteins indicating Pdx1-Pdx2 complex formation, lane A corresponds to Pdx1, lane B to Pdx2, lane C to the Pdx1–Pdx2 complex sample applied on SEC column, and D shows the SEC-purified Pdx complex; (**c**) Native PAGE analysis showing the unreduced form of the Pdx proteins: lane A: Pdx1 dodecamer; lane B: Pdx2 in different oligomeric states; lane C: a Pdx1 and Pdx2 sample mixed in 1:1 molar ratio applied to SEC column; lane D: purified Pdx complex with approximately 700 kDa [96]. **(d)** SEC elution profile of *S.aureus* Pdx complex and **(e)** PAGE evaluation of the SEC eluted peaks samples. Peak-1 corresponds to the oligomeric Pdx1, indicating low binding affinity to Pdx2, with Pdx2 remaining unreacted in peak-3. Peak-2, eluted at approx. 19 mL, is the *Sa*Pdx complex eluted from the SEC applied P1:P2 (Pdx1:Pdx2) sample.

4.3. DLS Investigations of *P.vivax* Pdx proteins

To characterize the stability and size distribution of the SEC eluted Pdx proteins in solution, the dynamic light scattering technique was applied. Dynamic light scattering is a useful technique to investigate the size distribution and dispersity of nanoparticles or

macromolecular complexes in solution. This method works on the principle of laser light (red-light) tracking the Brownian motion of particles in solution. The fluctuations in the laser scattering due to the particle's relative movement are correlated to the hydrodynamic radius (R_H) of the scattering particles, which are then used to calculate their molecular weight estimation in homogenous solutions [128].

For the monodisperse *P.viva*x Pdx1 solution, DLS revealed a hydrodynamic radius of (R_H) of 7.3 ± 0.9 nm **(Figure 18a)**, indicating its dodecameric state in solution with a reported Pdx1 diameter from homologous species of approx. 120 Å [69]. For the Pdx complex, a hydrodynamic radius of 9.7 ± 0.2 nm was observed **(Figure 18b)**. For Pdx2, DLS experiments revealed a concentration and time-dependent oligomerization tendency, with resulting R_H values of 2.8 ± 0.4 nm at 1.0 $mgmL^{-1}$ concentration, corresponding to approximately 35 kDa **(Figure 18c)** and 3.5 ± 0.4 nm at 1.6 $mgmL^{-1}$ and 4.2 ± 0.4 nm at 2.5 $mgmL^{-1}$, corresponding to approximately 60 kDa and 90 kDa respectively. The R_H values are also mentioned on their respective radius distribution profiles **(Figure 18d)**.

Interestingly mixing the oligomeric Pdx2 suspension, displaying multiple radius distribution values via DLS **(Figure 18e)**, with dodecameric monodispersed Pdx1. The mixing results in a monodispersed Pdx complex with a hydrodynamic radius of 10.9 ± 1.4 nm and a molecular mass of around 720 kDa **(Figure 18f)**. The observation indicates that Pdx2 is showing reversible oligomerization and comes out of its oligomeric state upon binding with its partner Pdx1, but this will require further experiments to confirm.

4.3.1. Reversible oligomerization investigation of *Pv*Pdx2 by DLS

Hence, *Pv*Pdx2 tends to form oligomers over time and concentration and further, after mixing with dodecameric Pdx1 it forms the Pdx1-Pdx2 complex. The complex forms a monodisperse suspension and a detailed discussion is provided in **figure 18f** legend. The reversible oligomerization behavior of *Pv*Pdx2 observed in preliminary DLS results required more confirmatory investigations by applying time-resolved DLS. In time-resolved DLS studies, Pdx complex formation was observed by utilizing the homogenous solutions of dodecameric Pdx1 and a Pdx2 (monomeric and oligomeric) solution in independent experiments. The obtained results are shown in **figure 19 a,b**. In the displayed results, the heat plots in the center represent the respective size

distribution over time with the abundance of radii (number of particles scattering the laser light) is color-coded from blue (low) to red (high).

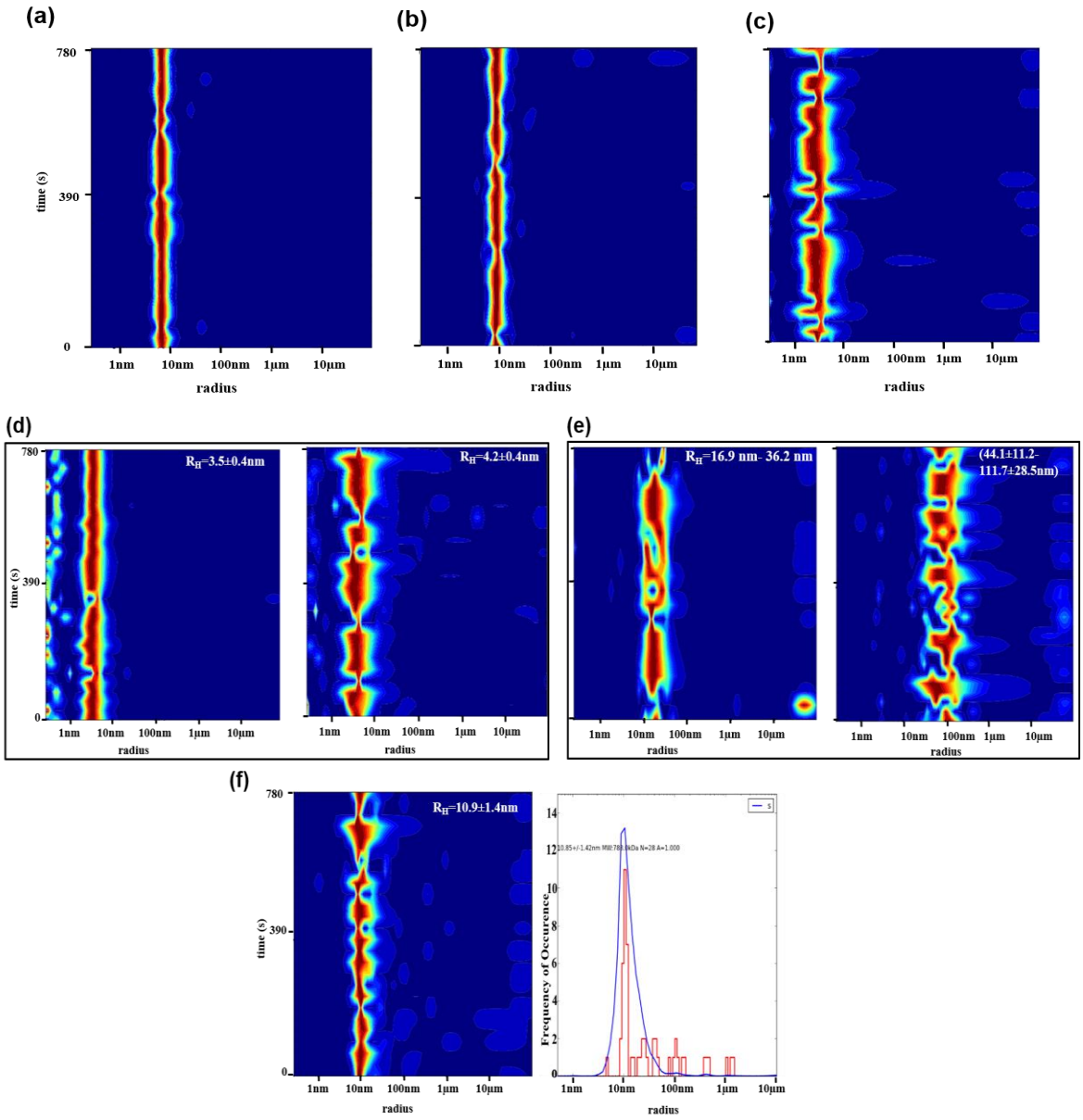


Figure-18. Dynamic light scattering (DLS) analysis of *P.vivax* Pdx proteins.
(a) Dodecameric Pdx1 (R_H = 7.3 ± 0.9 nm), **(b)** monodispersed saturated Pdx1–Pdx2 complex (R_H = 9.7 ± 0.2 nm), and **(c)** monomeric Pdx2 (R_H = 2.8 ± 0.4 nm). **(d)** Dynamic light scattering investigation of multimeric and oligomeric Pdx2, measured at a different time and concentrations, showing different oligomeric states, obvious from the radius distribution values (R_H of 3.5 at 1.6 $mgmL^{-1}$, R_H of 4.2 at 2.5 $mg\ mL^{-1}$, **(e)** R_H of 16.9-36.2 nm at a concentration of 3 $mg\ mL^{-1}$ of 10 days old sample and 44.1-111.7 at 9.5 $mg\ mL^{-1}$ of the freshly purified sample **(f)** Oligomeric Pdx2 after mixing with dodecameric Pdx1 in a 1:1 volume ratio and centrifuging at 10,000 g for 10-15 minutes, resulting in a Pdx complex with R_H = 10.9 ± 1.4 nm, shown by the radius distribution profile and histogram. Centrifugation is supporting the complex formation/interaction between oligomeric Pdx2 and dodecameric Pdx1, resulting mainly in a monodisperse solution [96].

In the first experiment, the dodecameric Pdx1 and monomeric Pdx2 solutions were mixed in a 1:1 molar ratio. Soon after mixing, the interactions between the two proteins were gradually established. Which results in a monodispersed Pdx complex at a concentration of 5 mgmL^{-1} with R_H = 10.4 nm **(Figure 19a)**.

In another parallel experiment, where a non-homogeneous solution of oligomeric Pdx2 was allowed to mix with the homogenous dodecameric Pdx1 solution. Mixing results in a Pdx complex formation with slightly higher R_H value distribution over time, compared to the size information obtained in the first experiment. However, the obtained suspension does not entirely show a monodispersed Pdx complex; presumably, some high-molecular-weight Pdx2 multimers remained in solution, even 12 hours after the mixing and subsequent measurement **(Figure 19b)**. Interestingly, a monodispersed Pdx complex solution was obtained after centrifuging suspensions of homogenous Pdx1 and non-homogeneous oligomeric Pdx2 at 10,000× *g* for 10–15 min, as shown in **figure 18f.**

The individual snapshots over the time course of the experiment, to investigate the reversible oligomerization behavior of *Pv*Pdx2, are documented and recorded in **figure 20a, b**. Figure **20a** includes the snapshots of the time-resolved DLS data of homogenous Pdx2 interaction with dodecameric Pdx1 over time. The resulting homogenous solution of the Pdx1-Pdx2 complex indicates the reaction dynamics of Pdx1 and Pdx2 molecules over the time course. Whereas figure **20b** represents the time-resolved DLS snapshots, recorded after mixing the non-homogeneous oligomeric Pdx2 and dodecameric Pdx1 solutions. The resulting data show the gradual formation of the Pdx1-Pdx2 complex over time with some remaining polydispersity. The time-resolved DLS data demonstrate that Pdx2 interaction with a Pdx1 dodecamer is independent of the solution state of Pdx2.

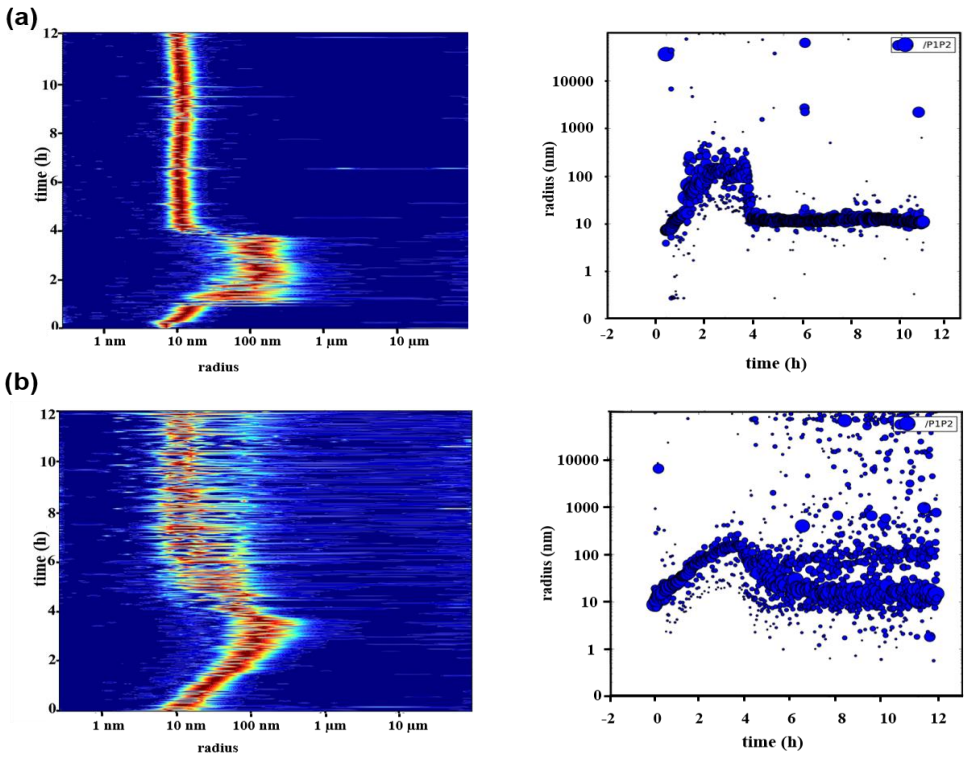


Figure-19. Reversible oligomerization of *Pv*Pdx2 by time-resolved DLS investigation.
Figures on the left in red indicate the particle size distribution, with the y-axis corresponding to the time course of the experiment and the horizontal axis showing the radius distribution of the particles scattering laser light. For figures on the right side, the axes are inverse. **(a)** The particle size distribution of dodecameric Pdx1 mixed with monomeric Pdx2 was measured over 12 h applying DLS. First, dodecameric Pdx1 was measured for 25 min after a solution of Pdx2 was added. After mixing Pdx1 and Pdx2, the predominant particle radii distribution shifted towards larger radii of approximately 100–300 nm. The high-molecular-weight particles (R_H > 100 nm) remained for two to three hours. Subsequently, the particle radii shifted to approximately 10–14 nm, corresponding to the radius of the Pdx complex; **(b)** DLS measurements of a suspension containing dodecameric Pdx1 and oligomeric Pdx2 recorded over time. After 15 initial measurements of dodecameric Pdx1, oligomeric Pdx2 was added and the radius profile shifted towards higher radius values, i.e., oligomeric Pdx2 interacted with Pdx1 to form the Pdx complex. After 4–5 h, the equilibrium shifted towards a radius corresponding to the Pdx complex radius with some remaining polydispersity. As seen from the radius profile, some large oligomeric particles, presumably corresponding to Pdx2 oligomers, remained in the solution [96].

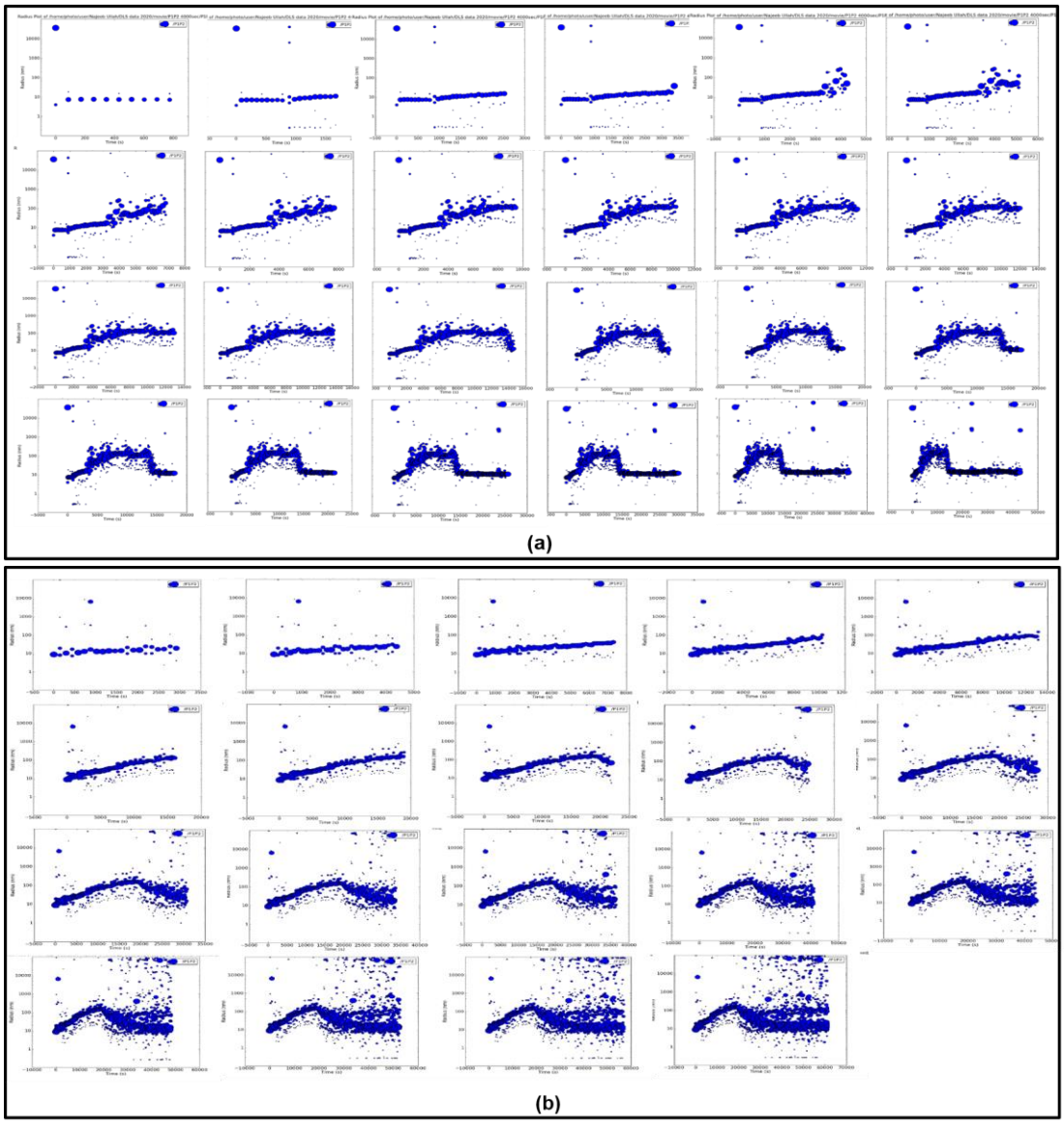


Figure-20. Long-term DLS measurements following the Pdx complex formation.

(a) Long-term DLS measurements showing the particle radius evolution of Pdx complex formation after mixing dodecameric Pdx1 and monomeric Pdx2 in a ratio of 1:1 in DLS cuvette with concentrations of 5 mgmL^{-1}. The measurement was started with a solution of the Pdx1 dodecamer for 15 initial measurements (30 seconds each and 60 seconds pause in between), then Pdx2 was added. The addition results in a direct increase in the hydrodynamic radius distribution, and after 4-5 hours, the particles gain a constant radius of approx. 10 nm (hydrodynamic radius of the Pdx complex). **(b)** Time-resolved DLS measurements of Pdx complex showing first R_H values obtained for dodecameric Pdx1 at a concentration of 5 mgmL^{-1}. After 15 initial measurements, 30 seconds each with 60 seconds pause time in-between measurements, Pdx2 was added, having a final Pdx1:Pdx2 volume ratio of 1:1. The hydrodynamic radii of the particles in the suspension increase slowly and after 4 to 5 hours the particles gained almost a constant radius distribution of approx. 11-15 nm (hydrodynamic radius of the Pdx complex) with some remaining polydispersity.

4.4. Nanoparticle Tracking investigation of *P.vivax* Pdx2 oligomerization

NTA is a technique that accurately analyzes the size distribution of particles in a monodisperse as well as the polydisperse sample solution. In NTA, the sample visualization and individual particle tracking in the mixture are its characteristics that enable a thorough size distribution analysis of individual particles in the suspension. Therefore, the NTA technique was utilized to obtain a clear picture of the reversible oligomerization tendency of *Pv*Pdx2 in the solution. In the experimental setup, the suspension containing homogenous dodecameric Pdx1 and non-homogenous oligomeric Pdx2 were allowed to mix and then observed over time for tracking the particles. After mixing both the solutions, the analysis confirms the disappearance of oligomeric Pdx2 peaks from its non-homogenous solution, and new peaks appear at the Pdx complex diameter range, as shown in **figure 21**. The resultant Pdx1-Pdx2 complex particles showed a diameter distribution of 25 to 35 nm **(Figure 21)**. Further, the decrease of the initial predominant 120-230 nm diameter distribution (red line), corresponding to the oligomeric Pdx2 particles, indicates the dissolving of oligomeric Pdx2 species and uniform Pdx1-Pdx2 complex formation. This reversible oligomerization of *Pv*Pdx2 observed by NTA investigations also support the interpretation of the time-resolved DLS data, which are shown in **Figures 19b and 20b**.

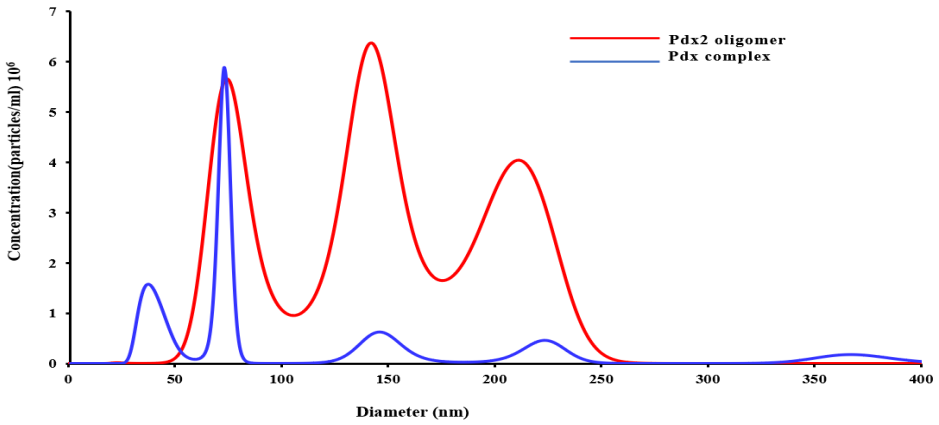


Figure-21. Nanoparticle tracking analysis of *P.vivax* Pdx complex formation.

NTA data showing Pdx complex formation from dodecameric Pdx1 and oligomeric Pdx2. First, the oligomeric Pdx2 (red line) was analyzed and recorded for nanoparticle tracking and then the dodecameric Pdx1 was added and allowed to mix. This results in Pdx1-Pdx2 complex formation (blue line) of diameter around 25-35 nm, which is very close to hydrodynamic radii (R_H) value of 11-15 nm for the Pdx complex, obtained from DLS measurements shown in **figure 19b and 20b**.

4.5. DLS investigation of *S.aureus* Pdx proteins

For monodisperse *S.aureus* Pdx1 solutions DLS studies showed a hydrodynamic radius (R_H) of 6.2 ± 0.3 nm at different concentrations **(Figure 22a)** with a corresponding molecular weight of 218 kDa, indicating a hexameric state in solution. For Pdx2, the DLS experiments revealed an R_H value of 2.7 ± 0.9 nm at concentrations ranging from 1-5 $mgmL^{-1}$, corresponding to approximately 31 kDa, indicating a monomeric state in solution **(Figure 22b)**. The hydrodynamic radius results indicate a slight increase in R_H value at increasing concentrations of Pdx2, but overall, the monomeric state was dominant. For the Pdx complex size and oligomeric state visualization, DLS data showed a hydrodynamic radius of 6.6 ± 0.2 nm **(Figure 22c)**. The radius value indicates either binding of a few Pdx2 molecules to *Sa*Pdx1 or some hexamer disintegration and structural rearrangement occurring in Pdx1 upon binding to Pdx2. For Pdx1 and Pdx complex, DLS measurements were performed at concentrations from 1-15 $mgmL^{-1}$, revealing almost consistent hydrodynamic radius values at all concentrations.

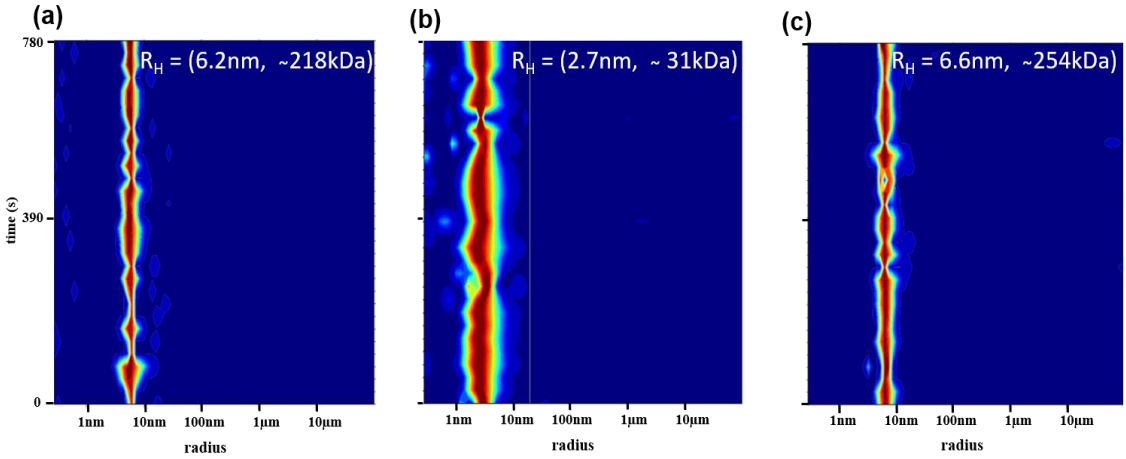


Figure-22. Dynamic light scattering (DLS) analysis of *S.aureus* Pdx proteins.
(**a**) Hexameric Pdx1 (R_H = 6.2 nm) with approximate molecular mass of 218 kDa (**b**) Pdx2 (R_H = 2.7 nm) indicating its monomeric form with molecular weight of 31 kDa, and (**c**) monodispersed Pdx1–Pdx2 complex (R_H = 6.6 nm) with molecular mass value of 254 kDa. The increase in hydrodynamic radius and molecular mass of the Pdx1-Pdx2 complex indicates, either a few Pdx2 molecules attached to Pdx1 hexamers or disintegration of Pdx1 hexamers into smaller multimers upon the formation of the Pdx complex.

4.6. Circular Dichroism (CD) Spectroscopy investigation of Pdx proteins

To evaluate the proper folding and secondary structure information of the purified Pdx proteins, CD spectroscopy was applied. For secondary structure investigation of *S.aureus* Pdx proteins, Pdx1 showed 46% of α-helices and 13% of beta-sheets. For

*Sa*Pdx2 α-helices content of 32% and beta-sheets content of 26% were observed. Whereas, for *P.vivax* Pdx1 52% α-helix and 20% β-sheet content, 36%, 34% for *Pv*Pdx2, and 54%, 25% for the complex was observed. The secondary structure content data show mean RMS (root mean square) values of 8.4, 6.2, and 6.5, respectively, after evaluating the recorded data by the JASCO Spectra Manager software suite [127]. The obtained results are displayed in detail in **figure 23** and table 13. The obtained CD data indicate two discrete minima at a wavelength of 207 nm and 222 nm, which are characteristic of a predominant α-helical secondary structure content [125, 163]. Further, the resultant β-sheet and random coil values are also given in table 13.

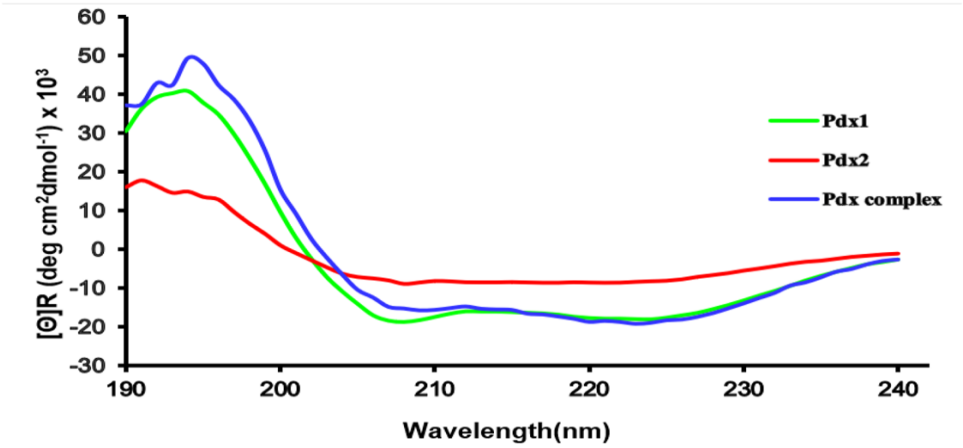


Figure-23. Circular Dichroism (CD) spectroscopy data of *P.vivax* Pdx proteins.
CD results showing secondary structure contents of *P.vivax* Pdx1 (green), the complex (blue) with α-helices and β-sheets as predominant secondary structure elements, and Pdx2 (red) with α-helices, β-sheets, and some random coils. Subsequently, the resultant experimental CD data obtained for Pdx proteins were also compared with secondary structure content predicting online servers, i.e., SOMPA, DSSP, and Predict Protein. The predicted data are consistent with the experimental data [96].

The CD data and the corresponding secondary structure contents were compared to the secondary structure content of the homologous reported structures. *P.vivax* Pdx1 CD data were compared to the reported *P.berghei* Pdx1 structure (Pdb code: 4ADU) [69] and *Pv*Pdx2 was compared with the *Pf*Pdx2 crystal structure (Pdb code: 2ABW) [103]. Further, *P.vivax* and *S.aureus* Pdx1, Pdx2 and Pdx complex secondary structure contents were calculated by applying different online servers, predicting the secondary structure contents i.e., SOPMA (self-optimized prediction method with alignment), DSSP (define the secondary structure of proteins), and Predict Protein. The resulting values are described in table 13 and are compared to the experimentally determined

values. The calculated data indicate very close consistency to the reported, as well as to the online prediction servers data.

Table-13: *P.vivax* and *S.aureus* Pdx proteins secondary structure content comparison obtained from CD data, predicted from online servers, and compared to homologous structures.

Protein	Experimental/Prediction Server/reported	α-helices (%)	β-sheets (%)	Coils (%)	Turns (%)
***Pv*Pdx1**	*Pv*Pdx1 recorded	52	20	22	6
	SOPMA	52	13	27	6
	DSSP	46	14	39	-
	Predict Protein	49	8	43	-
***P. berghei* Pdx1**	Pdb code: 4adt	43-52	12-13	-	-
***Sa*Pdx1**	*Sa*Pdx1 data recorded	46	13	24	12
	SOPMA	47	14	20	19
	DSSP	36	16	46	-
	Predict Protein	41	7	52	-
***Pv*Pdx2**	*Pv*Pdx2 recorded	36	34	30	0
	SOPMA	33	21	40	6
	DSSP	29	32	39	-
	Predict Protein	29	27	44	-
***P. falciparum* Pdx2**	Pdb code: 2abw	27-34	30	-	-
***Sa*Pdx2**	*Sa*Pdx2 recorded	32	26	18	20
	SOPMA	43	22	28	7
	DSSP	25	27	47	-
	Predict Protein	24	37	38	-
***Pv*Pdx complex**	*Pv*Pdx complex recorded	54	25	21	0
	SOPMA	40	18	33	9
	Predict Protein	36	18	46	-
***Sa*Pdx complex**	SOPMA	49	16	27	8
	Predict Protein	35	24	39	-

4.7. Mass photometry investigations

Mass Photometry (MP) works on the principle of quantifying individual molecules by illuminating and thereby revealing the interface between the sample and cover glass, recording the reflectivity changes. The reflectivity changes happen by the change in the local refractive index when an adhering biomolecule replaces water/buffer molecules [147]. Further optimization and subsequent improvement in the image contrast fields [164] enable very accurate quantification of these reflectivity changes of individual biomolecules and lead to mass measurements with good resolution and precision [150]. A key advantage of mass photometry over other existing mass-determining biophysical and bioanalytical solution-based approaches is its ability to distinguish directly between different species contributing to a polydisperse system. It is very sensitive in

quantification and accurate in determining the relative abundances of different biomolecules and their complexes in a polydisperse mixture [165].

4.7.1. Oligomerization studies of *P.vivax* Pdx1

*Pv*Pdx1 SEC eluted peak was investigated for concentration and time-dependent oligomerization by mass photometry. The results indicate that dodecameric *Pv*Pdx1 is independent of both concentration and time-based oligomerization after measured at different concentrations i.e. 10, 100 nM, and 1µM, as shown in **figure 24**. The obtained data are contrary to its plasmodial homologs showing oligomerization [69].

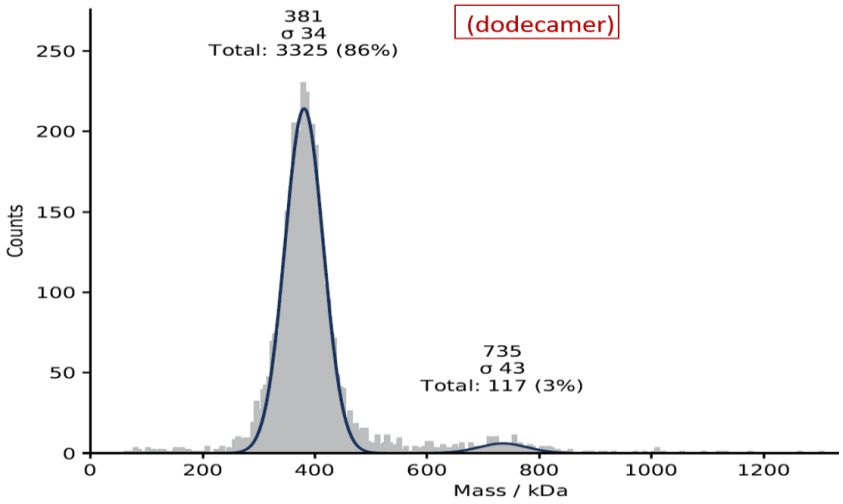


Figure-24. Mass photometry studies of *P.vivax* Pdx1.

MP investigation of *Pv*Pdx1 results showing 381 kDa molecular weight particles indicating the dodecameric state of Pdx1 irrespective of time and concentration. There is a small peak at 735 kDa corresponding to a 3% contribution to the mixture molecules. Since no higher-order oligomers were found by other biophysical methods, therefore it was considered as an artifact.

4.7.2. Dynamic oligomerization behaviour of *Sa*Pdx1

Fresh size exclusion chromatography purified *S.aureus* Pdx1 was investigated for oligomerization from the eluted peak at different concentrations. Interestingly, *Sa*Pdx1 was observed in trimer-hexamer equilibrium with an equilibrium position more towards trimer at the lower concentration of approx. 10^9 particles in solution. The equilibrium position shifts more towards the hexameric form of Pdx1 with an increase in the concentration until it is completely in hexamer form at a concentration of 1 µM (with approximately 10^{11} particles). The concentration-dependent oligomerization of *Sa*Pdx1 is shown in **figure 25**.

This unique observation of *Sa*Pdx1, which is not reported before for Pdx1 homologs, will entail further experimental investigations. In supportive evidence of the low molecular weight *Sa*Pdx1 species, there are reports from *P.falciparum* truncated Pdx1, which exist in low molecular weight species. *Pf*Pdx1 was found predominantly as a monomer with some dimer and trimer species also present in the solution after truncation at the C-terminus from 270-301. The monomeric Pdx1 was catalytically inactive but able to activate and form the Pdx complex with Pdx2 [166]. This also indicates that the Pdx1 hexameric or dodecamer state is a pre-requisite for its enzymatically active form.

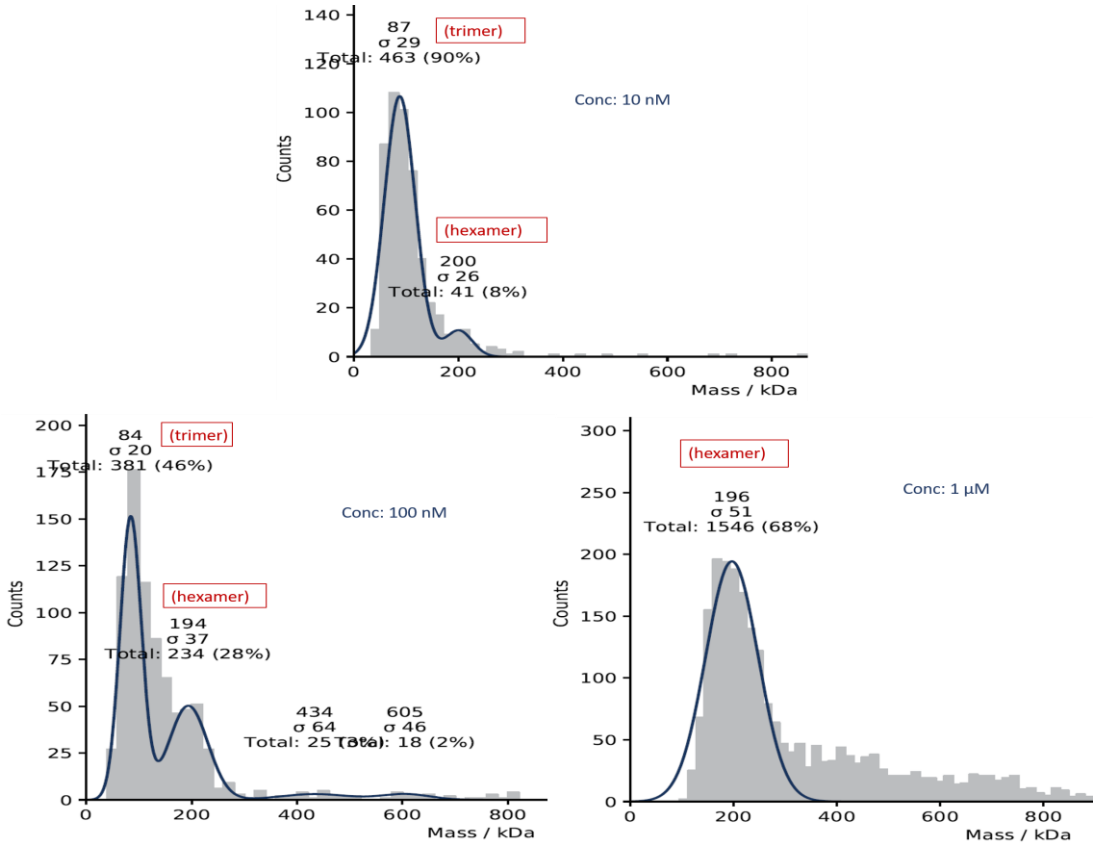


Figure-25. Concentration-dependent mass photometry investigation of *S.aureus* Pdx1.
The figure shows that at low concentration *Sa*Pdx1 is in trimer-hexamer equilibrium with the equilibrium position more towards the trimeric form and shift towards the high molecular weight with an increase in concentration. *Sa*Pdx1 molecules appeared in hexamer at a concentration of 1 µM and above, the concertation used in other biophysical and bioanalytical investigations.

4.8. SAXS (small-angle X-ray scattering) structural investigation

To further analyze the dynamicity, oligomerization, and visualize the 3D molecular structure of the Pdx proteins and their complex in solution, X-rays solution scattering (SAXS) investigations were performed. Here, monodispersed solutions of Pdx1, Pdx2, and Pdx1-Pdx2 complex were subjected to SAXS experiments and deduced their low-resolution 3D structural information.

4.8.1. SAXS analysis of *P.vivax* Pdx proteins

From X-ray solution scattering data for *Pv*Pdx proteins, the scattering intensities information of the Pdx proteins were obtained and their respective averaged scattering intensity profiles are displayed in **figure 26a–c**. Guinier analysis using the Guinier approximation and radius of gyration (R_G) for Pdx1, Pdx2, and their complex was determined by ATSAS program AUTOGNOM [133]. The obtained R_G values are summarized in table 14.

Interatomic distance distributions function *P(r)* generated from GNOM, is used to determine the quality of the sample as well as a general description of the overall structure of the protein [167]. It provides accurate real space approximations and can be used to generate low-resolution 3D models. The *P(r)* function indicates a maximum diameter (D_{max}) of 15 nm and a nearly globular shape for *Pv*Pdx1, a D_{max} of 11.3 nm, and a rod-shaped structure for *Pv*Pdx2. Whereas a D_{max} of 22.2 nm for the Pdx complex with a more spherical but slightly extended structure was obtained from the scattering data (**Figure 26d** and Table 14). Kratky plots ($I_{(S)}S^2$ versus S) obtained from the scattering data were used to verify the flexibility of the proteins. Based on the Kratky plots of the scattering data, *Pv*Pdx1 and the Pdx complex represent rigid and compactly folded particles **(Figure 26e)**. In contrast, the Kratky plot of Pdx2 indicated significant intrinsic flexibility **(Figure 26e)** [168]. Porod–Debye plots ($I_{(S)}S^4$ versus S) of the scattering data of the three proteins showed plateaus, indicating that Pdx1 and the Pdx complex lack disordered regions, whereas Pdx2 is more flexible in solution **(Figure 26f)**. The shape factor ($\rho = R_G/R_H$) which is used to provide the shape information [135], gave a value of 0.69 for Pdx1 and 0.71 for the Pdx complex, indicating both as nearly spherical globular particles. Whereas Pdx2 exhibited values ranging from 0.931 to 1.18, indicative of a flexible ellipsoidal structure [169]. All SAXS experimental parameters for *P.vivax* Pdx proteins are summarized in table 14.

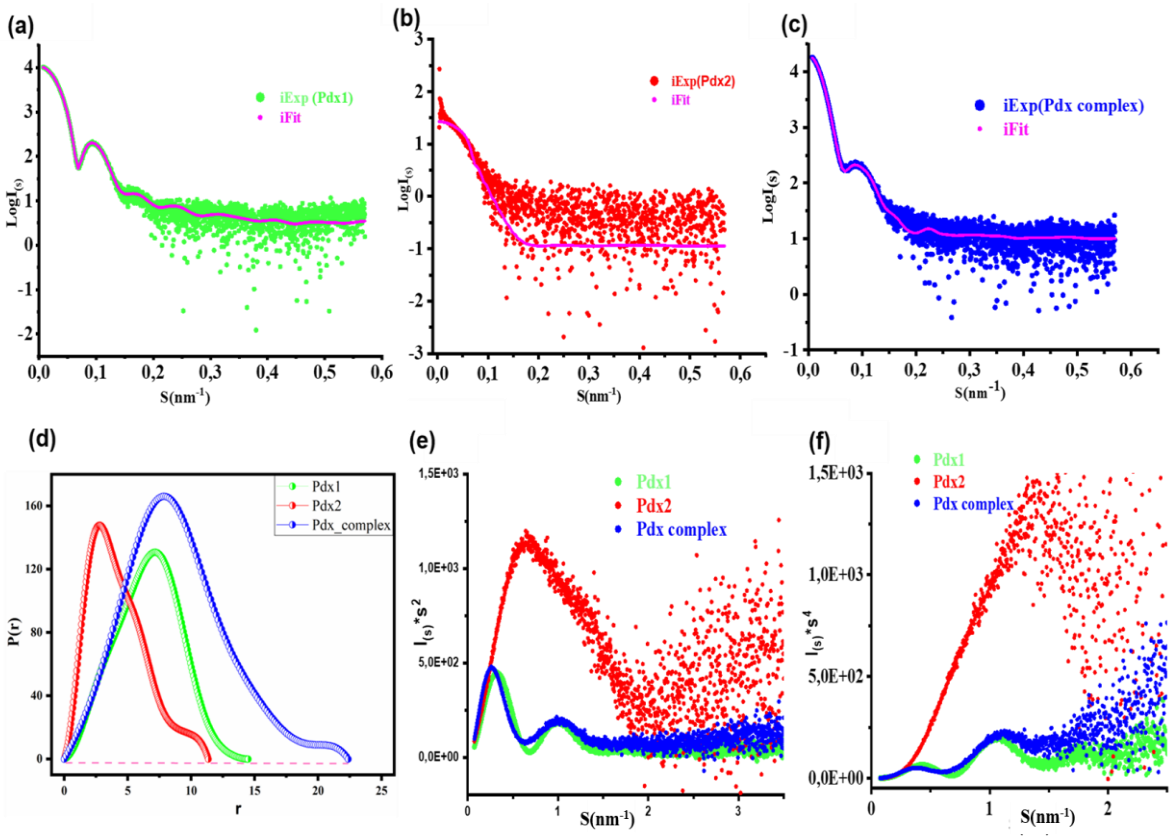


Figure-26. Processed solution scattering intensity pattern of *P.vivax* Pdx proteins.

(a) Pdx1 (green), (**b**) Pdx2 (red), and (**c**) the Pdx complex (blue) in arbitrary intensity units. Their fits with selected reported homolog structures are shown in magenta (Pdx1 (4ADU); Pdx2 (2ABW); Pdx complex: (4ADS). Data were processed and evaluated applying the CRYSOL and rigid body modeling program SASREF; **(d)** The distance distribution function *P(r)* indicates a maximum diameter (Dmax in nm) for Pdx1 (green), Pdx2 (red), and Pdx complex (blue) (**e**) Dimensionless Kratky plots indicate globular and compact structures for Pdx1 and the Pdx complex and a more flexible structure for Pdx2 (Pdx1 is shown in green, Pdx2 in red, and the Pdx complex in blue); (**f**) Porod–Debye plots of the SAXS data indicating properly folded proteins with no or least inter-particles interference and no aggregation, with the same color-coding used in (**e**) [96].

4.8.1.1. SAXS structures of *P.vivax* Pdx proteins

The scattering amplitudes for *P.vivax* Pdx proteins were obtained by processing the collected data using the program PRIMUS [132] from the ATSAS software suite. The scattering data were compared to the homologous crystal structures of *P.falciparum* and *P.berghei* with the pdb codes Pdx1, 4ADU; Pdx complex, 4ADS, and Pdx2, 2ABW **(Figure 26a–c)** using the program CRYSOL [140] and the rigid body modeling program SASREF [139]. The experimental scattering curves, when compared with the calculated scattering amplitudes, confirmed the structural similarities to the respective

homologous three-dimensional structures over a wide range of angles. Further, 3D models for Pdx proteins were constructed using online homology modeling approaches [106, 170], where Pdx1 was modeled as a dodecamer, Pdx2 in monomeric form, and the Pdx complex as 24-mer. A 3D model for Pdx1 was calculated based on *Pb*Pdx1 (pdb code: 4ADU) [69], for the Pdx complex based on the plasmodial chimeric complex (4ADS) [69] and for Pdx2 based on *P.falciparum* with pdb entry 2ABW [103].

The reported crystal structures of Pdx proteins from *P.falciparum* and *P.berghei* showed sequence identities of 88% and 85% for Pdx1, respectively, and 75% and 69% for Pdx2, respectively when compared to the *P.vivax* counterparts [171]. The calculated 3D models of Pdx1, Pdx2, and the Pdx complex were superimposed by applying the program Pymol [172] with the ab-initio-built GASBOR models [138] individually and fitted well with the dummy sphere models. Pymol superimpositions were performed for Pdx1 (dodecamer) with its ab-initio built model, and its fit indicated Pdx1 to be a dodecamer in solution. Further, the SAXSMoW data resulted in a molecular mass of 418.2 kDa for Pdx1 [173] (theoretical mass value for dodecameric Pdx1 is 417.3 kDa) indicating its dodecameric state in solution. For the *Pv*Pdx complex, a molecular mass estimation of the SAXS data indicates 735.8 kDa [173], very close to the theoretically calculated mass of 730.5 kDa.

The obtained rigid body SASREF [139] models were separately superimposed with the ab-initio models of Pdx1 **(Figure 27a)**, Pdx2 **(Figure 27b)**, and the Pdx complex **(Figure 27c)** respectively, indicating an appropriate fit for Pdx1 and the Pdx complex. Due to reversible oligomerization behavior and the variable oligomeric states of Pdx2, rigid body models were built and calculated in regard to monomeric, dimeric, and elongated trimeric forms. The dimeric rigid body model fit best when superimposed with the ab-initio models **(Figure 27b)**, considering an additional significant contribution of the trimeric Pdx2 to the scattering intensity pattern, as well as in the 3D models. A significant contribution of the trimer was also indicated by the SAXSMoW data, with a corresponding molecular mass value of approximately 67 kDa, using the Pdx2 scattering profile. The theoretically calculated mass of a Pdx2 monomer is 26.1 kDa. All the ab-initio and rigid model superpositions were performed by applying Pymol [172].

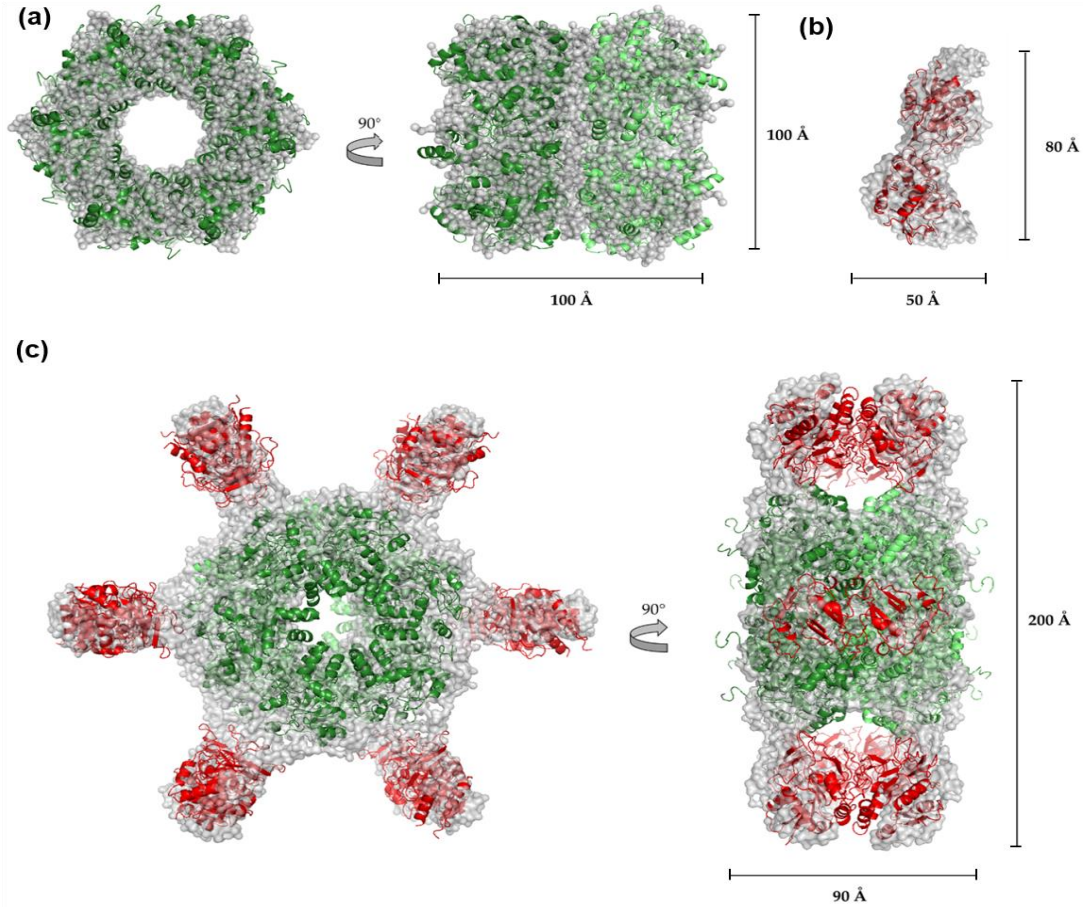


Figure-27. Ab-initio GASBOR and rigid body SASREF models of *P.vivax* Pdx proteins.

Ab-initio GASBOR models composed of light grey-color, chain-like dummy surface models superimposed with the rigid body SASREF models shown in cartoon representation. **(a)** Dodecameric Pdx1 rigid body model, i.e., one hexamer in green and one in lime color considering a P6 symmetry superimposed with its ab initio model; **(b)** P1 symmetry for dimeric Pdx2 in red fitted to its ab initio model; **(c)** ab initio model of the Pdx1–Pdx2 (12:12) complex possessing P6 symmetry with two subunits: two hexameric ring-shaped inner cores consisting of Pdx1 (green and lime color) and a total of 12 Pdx2 molecules (red) displaying a hexameric symmetry around the outside surface of both stacked Pdx1 hexamers, corresponding to the obtained ab initio model. The ab initio model is superimposed well with the rigid body Pdx complex [96].

Table 14. SAXS data collection and analysis parameters of *P.vivax* Pdx proteins [96].

Data collection parameters	**Pdx1**	**Pdx2**	**Pdx complex**
X-ray source	PETRA III; EMBL beamline P12		
Wavelength (nm)	0.124		
Detector distance (m)	3.1		
Temperature (K)	293	283	293
Structural parameters			
$I(0)$ (*P(r)* function)	10,720 ± 5.93	9161 ± 41.47	19380 ± 23.04
$I(0)$ (Guinier/Auto$R_{G)}$	10,727.20 ± 11.61	9307.96 ± 101.56	19,658.30 ± 28.99
$I(0)$ (SAXSMoW)	10,751.50	33.64	19,592.77
R_G (nm) (*P(r)* function)	5.00 ± 0.01	3.26 ± 0.02	6.83 ± 0.01
R_G (nm) (Guinier/AutoR_G)	5.05 ± 0.06	3.24 ± 1.36	6.95 ± 0.11
R_G (nm) (based on SAXSMoW)	5.07	3.52	6.89
qR_G limit (from Guinier/AutoR_G)	1.12	1.29	1.15
qR_G limit (from SAXSMoW)	1.29	1.30	1.29
D_{max} (nm) (from *P(r)* function)	15.0	11.3	22.2
Porod volume estimate (nm)	609.54	75.88	1733.18
$\rho = R_G/R_H$	0.69	0.93–1.18	0.71
Molecular mass determination (kDa)			
Molecular mass (SAXSMoW)	418.2	67.0	735.8
Molecular mass (amino acids sequence)	417.3	26.1	730.3
Oligomeric state	dodecamer	oligomer	24-mer

4.8.2. SAXS analysis of *S.aureus* Pdx proteins

To obtain low-resolution 3D structural information for *S.aureus* Pdx proteins, they were investigated by SAXS scattering analysis at different concentrations i.e. 2, 5, and 10 mgmL^{-1}. These different concentration SAXS measurements were planned to aid in the concentration-dependent oligomerization behavior of the Pdx proteins and particular *Sa*Pdx1, which was found in smaller multimeric forms at low concentration (nM to few µM) in mass photometry experiment **(Figure 25)**. SAXS scattering intensity information was obtained for *S.aureus* Pdx proteins and their respective averaged scattering intensity profiles, processed by utilizing the ATSAS software suite program PRIMUS [132], are displayed in **figure 28a–c**.

The Kratky plots ($I_{(S)}S^2$ versus S), obtained from the scattering data, represent *Sa*Pdx1 and *Sa*Pdx2 **(Figure 28d**) rigid and compactly folded proteins. In contrast to the

spherical and compact *P.vivax* Pdx complex, the Kratky plot data of the *Sa*Pdx complex indicate significant intrinsic flexibility and is more scattered **(Figure 28d**). Porod–Debye plots ($I_{(S)}S^4$ versus S) of *S.aureus* Pdx1 and Pdx2 obtained from their scattering data showed plateaus, indicating both in properly folded states. Whereas the Pdx complex with no plateau indicating its more flexible form and to some extent polydispersity in solution **(Figure 28e)**. The shape factor ($\rho = R_G/R_H$) [135] gave a value of 0.69 for *Sa*Pdx1 presenting its spherical globular shape and a value of 0.71 for the *Sa*Pdx2 showing to be nearly spherical particles. While the Pdx complex exhibited a value of 0.58, indicative of an open structure [169]. All SAXS experimental parameters of *Sa*Pdx investigated proteins are given in table 15.

Subsequently, the *Sa*Pdx proteins were also investigated for different states in the solution where all eluted peaks were investigated by SEC-SAXS at a concentration of 7 mg mL^{-1} each. The elution profiles **(Figure 29a,b,d)** with their respective interatomic distance distribution *P(r)* function are obtained from the X-ray scattering of the eluting molecules. The distance distribution *P(r)* functions are generated by the ATSAS program GNOM [133] and are shown in **figure 29c,e**. The *P(r)* indicates a maximum diameter (D_{max}) of 13.44 nm and a nearly globular spherical shape for Pdx1. *Sa*Pdx2 displayed a D_{max} of 5.71 nm, and a nearly spherical structure was deduced from the *P(r)* curve. Pdx complex provided a D_{max} value of 11.3 nm with the *P(r)* profile indicative of a more extended and multidomain structure **(Figure 29e)**.

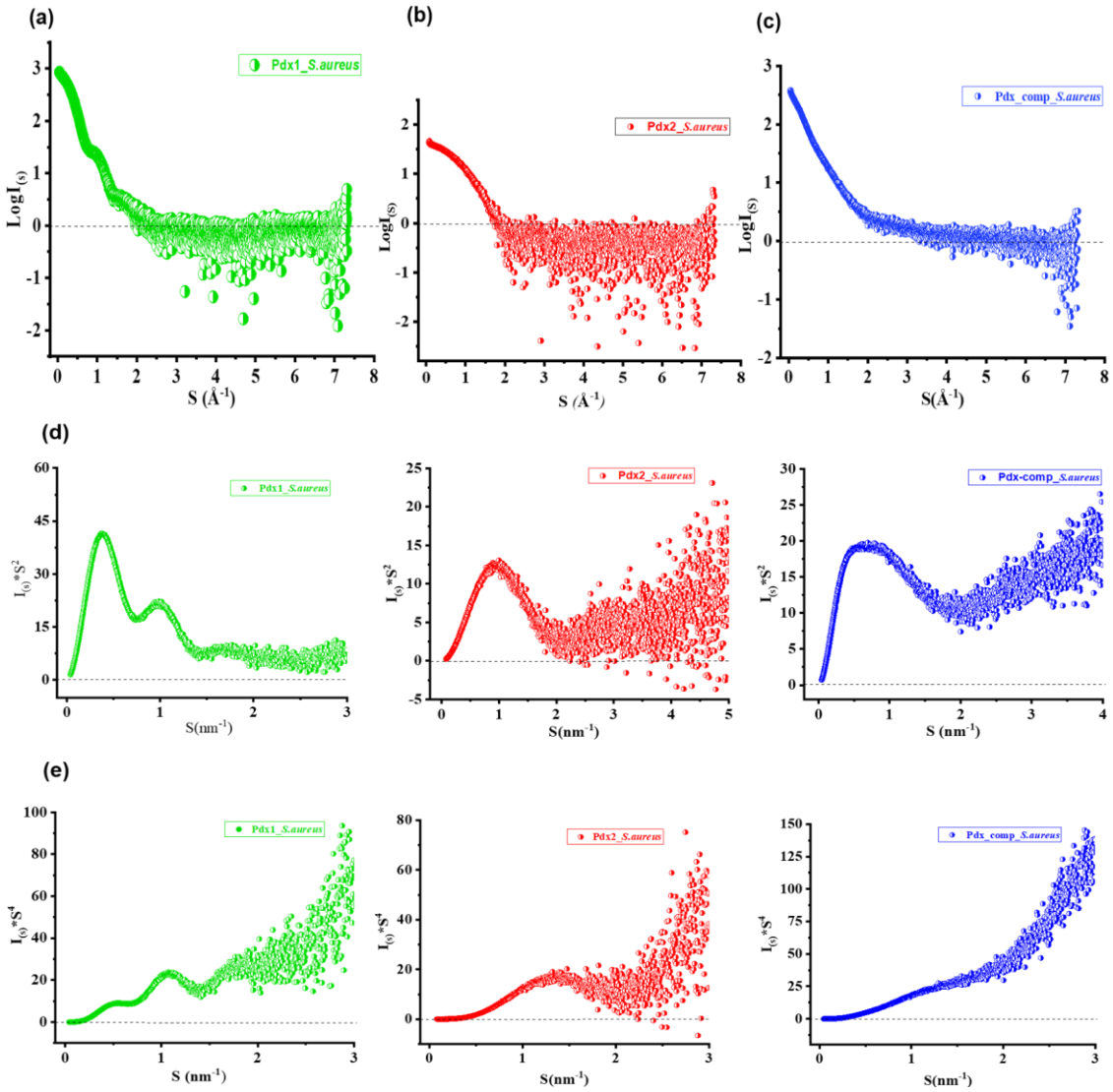


Figure-28. Processed solution scattering data of *S.aureus* Pdx proteins.

(a) Pdx1 (green), **(b)** Pdx2 (red), and **(c)** Pdx1-Pdx2 complex (blue) in arbitrary intensity units. Rigid body modeled 3D structures were calculated and constructed from these initial scattering intensities. A *Sa*Pdx complex scattering intensity profile was compared with its 3D predicted structure prepared by the I-tasser online server [106]. It was showing best fit with the tetrameric modeled structure model. Further, these scattering data were also compared to the reported data by applying the ATSAS programs CRYSOL [140] **(d)** The dimensionless Kratky plot indicates globular and compact structures for Pdx1 (green) and Pdx2 (red). Whereas, the Pdx complex demonstrates more flexibility and is partially unfolded (blue) **(e)** Porod–Debye plots generated from the scattering intensity data of *Sa*Pdx1 and Pdx2 proteins indicate properly folded proteins with no or least inter-particles interference, whereas, the Pdx complex shows some polydispersity**.**

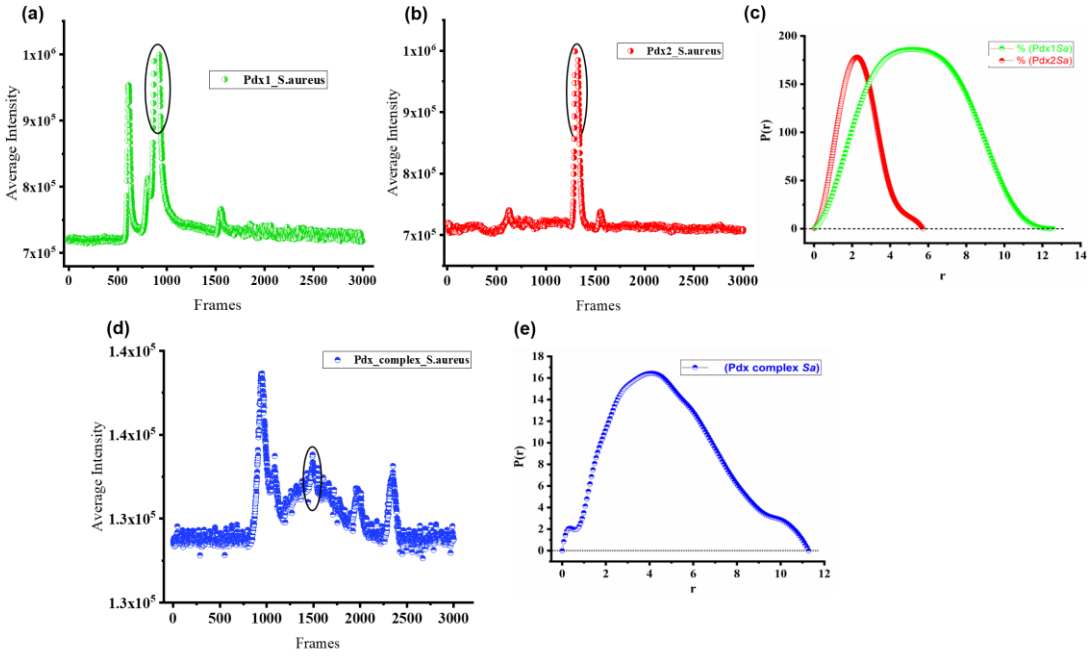


Figure-29. SEC-SAXS elution profile of *S.aureus* Pdx proteins with respective *P(r)* functions. **(a)** *Sa*Pdx1 eluted in the encircled peak with collected frames along X-axis and intensity along Y-axis **(b)** *Sa*Pdx2 eluted in one main monomeric peak **(d)** *Sa*Pdx1-Pdx2 complex is eluted at approx. 1500th frames range. **(c-e)** Interatomic distance distributions function *P(r)* graphs generated by ATSAS program GNOM, provide the preliminary structure information of the proteins and helps in 3D SAXS structures model construction. The *P(r)* indicates a maximum diameter (D_{max}) of 13.44 nm for Pdx1 (green), a Dmax of 5.71 nm for Pdx2 (red) **(c)**, and a D_{max} of 11.3 nm for the Pdx complex (blue), indicating a more extended and flexible structure **(e)**.

4.8.2.1. SAXS structures modelling of *S.aureus* Pdx proteins

To obtain the 3D models and structural information, the processed scattering amplitudes for *Sa*Pdx proteins were compared with the reported homologous *B.subtilis* Pdx proteins with Pdb codes for Pdx1: 2NV1 and for Pdx2: 2NV0 [90] utilizing the ATSAS programs CRYSOL [140] and SASREF [139]. Hence, *Bs*Pdx1 (2NV1) is a dodecamer, therefore *Saccharomyces cerevisiae* hexameric Pdx1 with Pdb codes of 3O06 [91] and 3FEM [99] were additionally used as reference models for comparison. Ab-initio built hexameric and dodecameric GASBOR models [138] were constructed for *Sa*Pdx1 and the dummy hexameric model was fitting the scattering profile. GASBOR generates a monomeric fitted model for *Sa*Pdx2 and the Pdx complex results in a tetrameric assembly of Pdx1 and Pdx2 (2 : 2).

Further, the 3D models for Pdx proteins were also constructed using the online homology modeling programs [106, 170], where *Sa*Pdx1 was modeled as a hexamer

and dodecamer and Pdx2 in the monomeric form. A dodecameric 3D model for *Sa*Pdx1 was calculated based on *Bs*Pdx1 (2NV1) [90], and a hexameric 3D model was calculated based on *Sc*Pdx1 (3O06 and 3FEM). For *Sa*Pdx2, a 3D model was constructed from *Bs*Pdx2 (2NV0) and for the Pdx1-Pdx2 complex, the symmetric hexameric (P_6) and tetrameric models were constructed applying the online modeling programs Swiss-model [170] and I-tasser [106]. The calculated 3D models for Pdx1, Pdx2, and the Pdx complex were superimposed by applying the program Pymol [172] with the ab-initio built GASBOR models [138] individually and fitted well with the dummy sphere models. Pymol superimpositions were performed for Pdx1 (hexamer) with its GASBOR model, and its fit indicates Pdx1 in a hexameric state in solution. Pdx2 superimposition indicates best fitting to its monomer form, whereas Pdx complex fits well with the tetrameric model upon superimposition.

Further, the online server SAXSMoW data indicate a molecular mass of approximately 208 kDa for *Sa*Pdx1 [173] which is close to its hexameric theoretical mass value of 192 kDa and indicating a hexameric form in solution. Both, batch and SEC-SAXS investigations reveal *Sa*Pdx1 in hexamer state **(Figure 30a)** with a molecular weight of 208 kDa. SAXS data support the DLS investigation for *Sa*Pdx1 (DLS section; **Figure 22**), which was suggesting a hexamer form with hydrodynamic radius value (R_H = 6.2 nm) and approximate molecular weight of 218 kDa. The mass photometry studies also indicate *Sa*Pdx1 in a dominant hexamer solution form at 1 µM (32 µg mL^{-1}) concentration **(Figure 25)**. For *Sa*Pdx2, a monomeric solution form was indicated by SAXSMoW, providing its molecular mass of 28 kDa calculated from the scattering data. For the Pdx complex, a molecular mass estimation of the SAXS data indicate 93 kDa [173], which is close to the theoretically calculated mass (106 kDa) of the tetrameric Pdx complex. The SAXSMoW data are provided in table 15.

The obtained rigid body SASREF [139] models were separately superimposed with the ab-initio constructed GASBOR models of the hexameric Pdx1 **(Figure 30a)**, monomeric Pdx2 **(Figure 30b)**, and the Pdx complex **(Figure 30c)** respectively, indicating appropriate fits. A Pdx complex SASREF model, with Pdx1 and Pdx2 in 1:1, was also superimposed with the model obtained from the online servers and fit best **(Figure 30d)**.

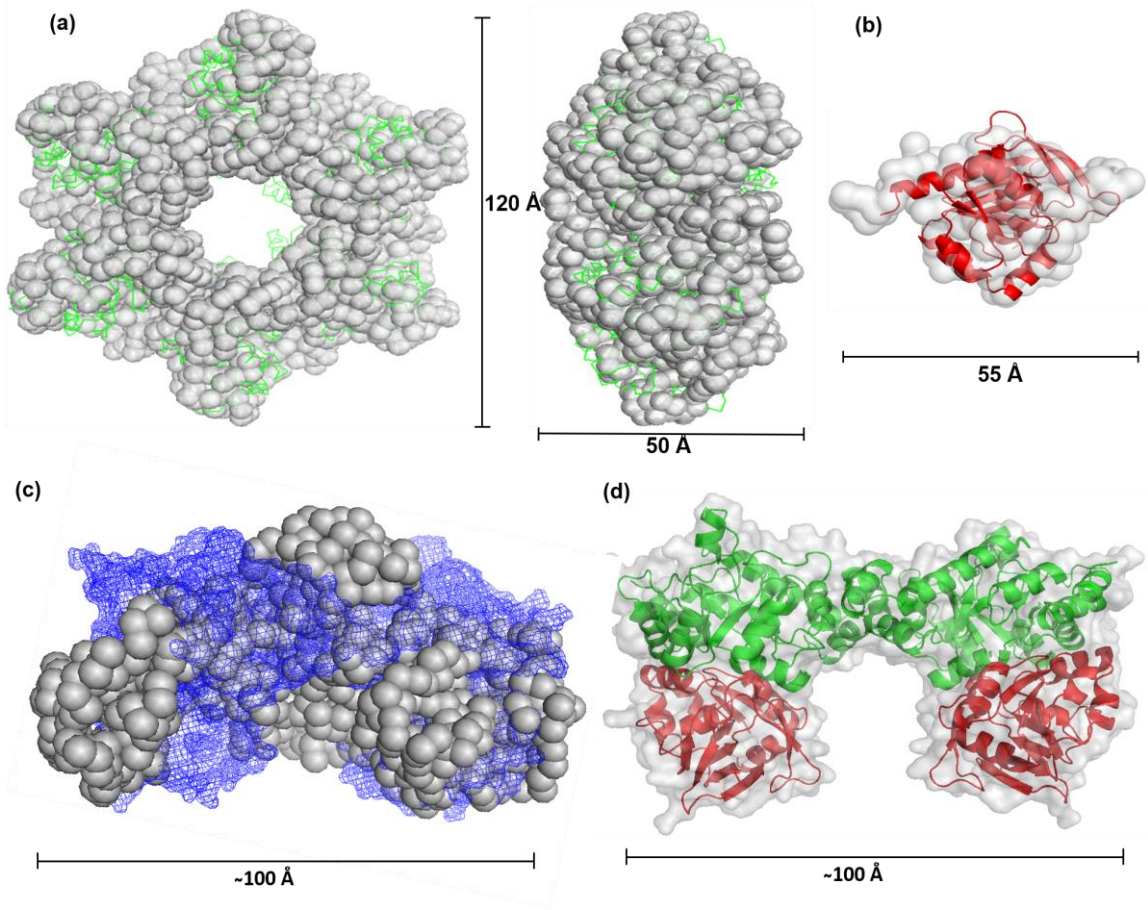


Figure-30. Ab-initio GASBOR and rigid body SASREF models of *S.aureus* Pdx proteins.
(a) Top and side view of *Sa*Pdx1 ab-initio GASBOR model composed of light grey-colored chain-like dummy spheres model, superimposed with the rigid body SASREF model shown in ribbon representation in green. Both modeling data support the hexameric solution state for *Sa*Pdx1. **(b)** Monomeric Pdx2 SASREF model in red fitted to its ab-initio GASBOR model. **(c)** Ab-initio model of the Pdx complex (light grey color beads) superimposed with the tetrameric modeled SASREF model shown in mesh (blue) and **(d)** the Pdx complex consisting of two Pdx1 (green color) and two Pdx2 molecules (red color), displaying a tetramer, corresponding to the SASREF model is superimposed well with its homology model shown in the light grey surface.

Table 15. SAXS data collection and analysis parameters of *S.aureus* Pdx proteins

Data collection parameters	Pdx1	Pdx2	Pdx complex
X-ray source	PETRA III; EMBL beamline P12		
Wavelength (nm)	0.124		
Detector distance (m)	3.1		
Temperature (K)	293	293	293
Structural parameters			
$I(0)$ ($P(r)$ function)	16,230 ± 7.2	56.37± 4.1	12.65 ± 0.09
$I(0)$ (Guinier/AutoR$_{G)}$	16,805.70 ± 27.4	56.48± 2.3	12.35
$I(0)$ (SAXSMoW)	16,772.03	53.58	12.49
R_G (nm) ($P(r)$ function)	4.44	1.86	3.77
R_G (nm) (Guinier/AutoR$_G$)	4.35	1.86	3.89
R_G (nm) (based on SAXSMoW)	4.29	1.85	2.80
qR$_G$ limit (from Guinier/AutoR$_G$)	1.16	1.30	1.30
qR$_G$ limit (from SAXSMoW)	1.29	1.29	1.29
D_{max} (nm) (from $P(r)$ function)	13.44	5.7	11.3
$\rho = R_G/R_H$	(4.29/6.2) = 0.69	(1.86/2.6) = 0.71	(3.82/6.6) = 0.58
Molecular mass determination (kDa)			
Mol. mass from SAXSMoW	208.0	28.0	93.0
Mol. mass from amino acids sequence	192.0	21.0	106.0
Solution oligomeric state	hexamer	monomer	tetramer

4.9. Electron Microscopy (EM) analysis

Negative staining electron microscopic observation of the particles offers the possibility for snapshots that elaborate direct information about their external shape, form, and structure [152]. Further, single-particle electron microscopy (EM) is then the required technique to determine the three-dimensional (3D) structure from those images collected in different orientations of individual molecules. In contrast to macromolecular crystallography, where one requires a high amount of protein and also molecules with more complexity and flexibility challenging to crystallize. The negative stain and single-particle EM require a minimal amount of sample with no condition for the sample to be in the crystalline form [153].

Considering the large oligomeric structure and complex assembly of Pdx1 and Pdx complex, they were investigated by negative stain electron microscopy (EM) to obtain and validate their solution insights. *Pv*Pdx1 structure was analyzed by negative-stain EM indicating dodecamer at around 30 µg mL^{-1}, and the resulting images are shown in **figure 31a,** with simple schematic representation in **figure 31b,c**. Negative stain EM images indicate Pdx1 in different orientations and thereby validating the characteristic

Pdx1 dodecameric assembly made up of two hexameric rings. Whereas *S.aureus* Pdx1 disintegrates during staining, and successful data acquisitions were not obtained.

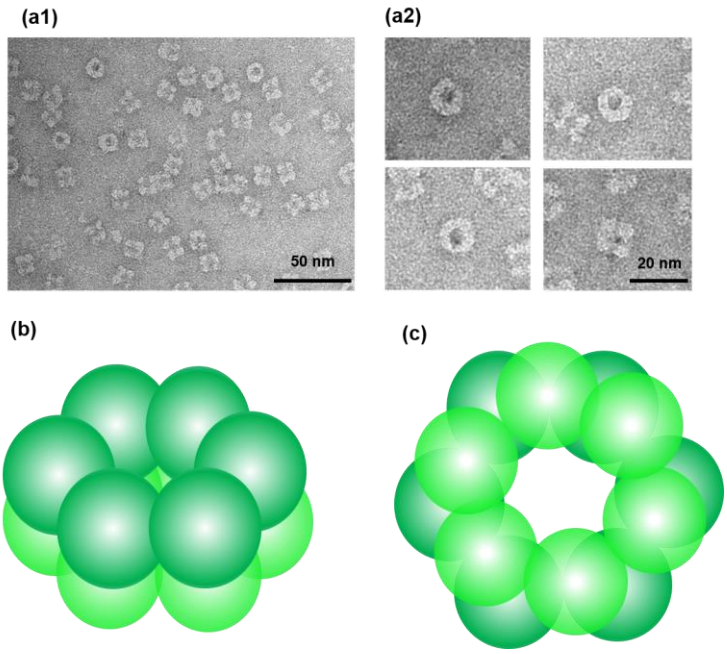


Figure-31. Transmission electron micrographs of negatively stained *P.vivax* Pdx1 protein. **(a1)** Dodecameric Pdx1, top view, and side view with random orientations of the particles. Images on the right in **(a2)** alongside a 20 nm scale bar, show a zoom-in of a representative class of averaged sample pictures shown on the left in (a1) [96]. **(b, c)** Schematic representation of the Pdx1 dodecameric assembly with a side view (left) and top view (right), respectively, with one hexamer in dark green and the other stacked hexamer in transparent light green circles.

Subsequently, the dynamic assembly of the Pdx complex was also analyzed by electron microscopy, showing a partial saturation of the spherical ring-shaped *Pv*Pdx1 dodecamer with *Pv*Pdx2 domains **(Figure 32a)**. This follows the reported assumption that the attachment of Pdx2 to the dodecameric core of Pdx1 takes place gradually [69]. Analysis of electron microscopy figures showed that Pdx1 and Pdx2 assemble transiently, and on average, a few Pdx2 subunits occupy Pdx1 dodecamers without preference for a distinct association pattern, illustrated in schematic **figure 32c**.

Further, the reversible oligomerization behavior of Pdx2 was also analyzed by EM to investigate and confirm the observed hydrodynamic radius shift recorded by DLS after mixing the dodecameric Pdx1 with monomeric Pdx2 in one experiment and with oligomeric Pdx2 in another experiment. The obtained EM micrographs confirmed the DLS data of the oligomeric Pdx2 binding to dodecameric Pdx1 and displayed that Pdx1 is complexed with Pdx2 (**Figure 32b**). As seen in Figure 32a, the monomeric Pdx2 binds to the dodecameric Pdx1; however, not all available vacancies were filled and the

complex is unsaturated. Whereas, binding of the oligomeric Pdx2 to the structured Pdx1 dodecamer result in a slightly asymmetric ring morphology of the Pdx complex particles, as shown in **Figure 32b**. The saturated Pdx complex particles shown in **Figure 32b** are also slightly larger, as illustrated in **figure 32d** than those shown for monomeric Pdx2 bound to Pdx1. However, the *S.aureus* Pdx1-Pdx2 complex was investigated for EM studies with different concentrations of the sample, but no successful results were obtained.

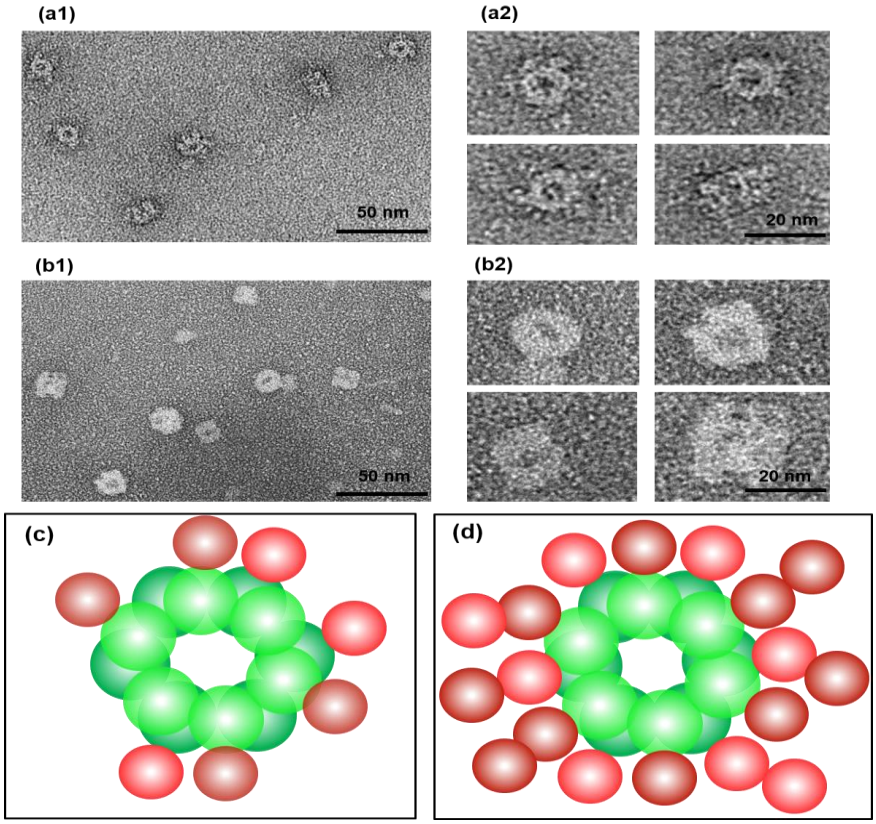


Figure-32. Electron micrographs and schematic representation of *P.vivax* PLP synthase, complexed from the dodecameric Pdx1 with monomeric and oligomeric Pdx2.
Images on the right in (a2), (b2) with a 20 nm scale bar, show a zoom-in of an averaged sample pictures shown on the left in (a1), (b1). **(a)** Dodecameric *Pv*Pdx1 in complex with monomeric Pdx2, a monomeric Pdx2 solution was reacted with Pdx1, for complex formation. Pdx complex particles were generally observed with partially bound Pdx2, whereas saturated (12:12) Pdx complex was rarely observed. **(b)** EM micrograph of dodecameric Pdx1 in complex with Pdx2 oligomers. The Pdx complex was predominantly saturated with Pdx2, showing a slightly larger dimension than those observed for the Pdx complexes shown in (a), demonstrating the larger hydrodynamic radius of the corresponding Pdx complex observed by DLS (Figure 19b and 20b) [96]. **(c,d)** EM visualization is represented by schematic diagram representation for Pdx complex; formed from monomeric Pdx2 **(c)** indicating unsaturation of the Pdx complex and oligomeric Pdx2 **(d)** mixed with Pdx1; indicating saturated Pdx complex with attached multimeric Pdx2, respectively.

4.10. Pdx proteins crystallization

4.10.1. *P.vivax* Pdx1 Crystallization

After optimizations of the purified *P.vivax* Pdx1 protein, required for crystallization and successful high throughput screening. It crystallizes at a concentration of 20 mg mL^{-1} by applying the hanging drop vapor diffusion method against a crystallization buffer of 0.1 M Tris (P^H = 8.0), 200 mM $MgCl_2$, 19% ethanol. Crystals were harvested and cryocooled in liquid nitrogen using crystallization buffer and an additional optimized 20% glycerol for cryo-protection. These flash-cryocooled crystals were then used for X-ray diffraction and data collection **(Figure 33a,b)**. Diffraction data were collected at beamline P11, PETRA-III (DESY, Hamburg). Details of the collected and processed data of *Pv*Pdx1 are summarized and presented in table 16.

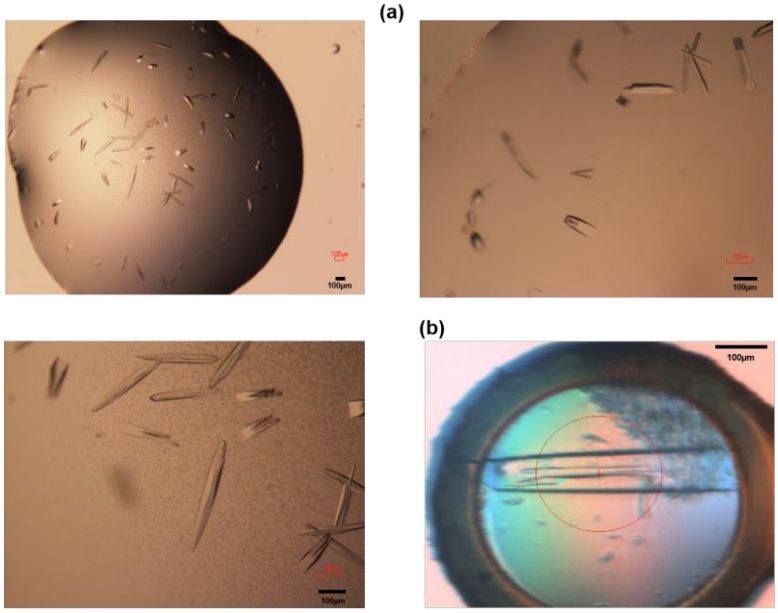


Figure-33. *P.vivax* Pdx1 crystals obtained after crystallization screens optimization
(a) *Pv*Pdx1 crystals of size approx. 50 x 600 µm before freezing in cryoprotectant **(b)** *Pv*Pdx1 crystal diffracted and data processed at 2.0 Å, the snapshot was taken right before the data collection at beamline P11 PETRA-III, DESY.

4.10.2. *P.vivax* Pdx1 dodecamer in the asymmetric unit (ASU)

The diffraction data of *Pv*Pdx1 were processed by using the software XDSAPP [120] to obtain the optimized reflection (mtz) file. Applying the Matthews probability calculator [174], molecular mass and accordingly, the number of subunits in the asymmetric unit (ASU) were determined, suggesting a dodecameric Pdx1 packed in the ASU with an occupied volume of 1188325 A^3. *P.vivax* Pdx1 crystallizes in space

group P 2_1 as dodecamer with a molecular weight of 408 kDa in the asymmetric unit, with a corresponding 57% solvent content. The *Pv*Pdx1 dodecamer packed in ASU electron density is shown in **figure 34a,b**. The structure figures of the model were prepared by UCSF Chimera [102].

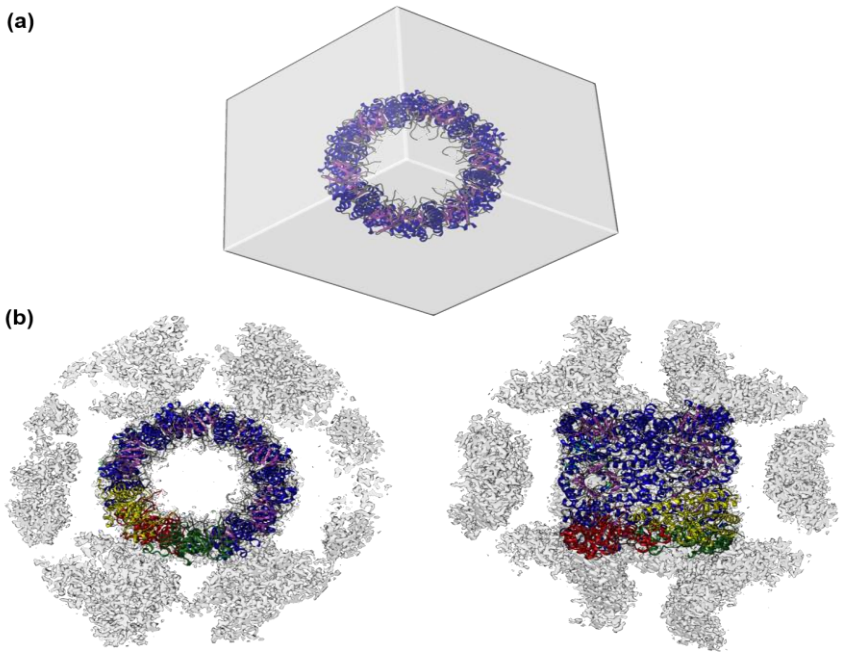


Figure-34. *P.vivax* Pdx1 dodecamer form was observed in the asymmetric unit.
(a) *Pv*Pdx1 dodecamer molecule of size ≈ 100 Å x 100 Å packed in ASU with cell dimensions in Å (119.969, 167.507, 120.208) **(b) left:** Pdx1 dodecamer top view with an electron density of the asymmetric unit indicating the internal empty tunnel in Pdx1 dodecamer; **right:** Pdx1 dodecamer side view with electron density in ASU, the spatial arrangement displays the independent dodecamers with spaces between each dodecamer, the figure is prepared by UCSF Chimera [102].

4.10.3. Oligomeric structure and unit cell arrangement of *Pv*Pdx1

Utilizing the ASU content information, *P.vivax* Pdx1 3D structure was determined by molecular replacement using *Plasmodium berghei* Pdx1 (Pdb code: 4ADT) with 85% sequence identity as a reference model [69] by applying the PHENIX program Phaser-MR [158]. The initial model was refined by rigid-body refinement using the program Phenix-refine [159] and model building was performed applying the program Coot [160]. The final refined model has R_{work} of 19.67 %, R_{free} of 22.16 %, and 1560 solvent molecules, as indicated in table 16. The final refined *Pv*Pdx1 dodecamer molecule packed in the unit cell is shown in **figure 35**. *Pv*Pdx1 crystallized into dodecamer, is forming a cylinder of approx. 100 Å tall and 100 Å in diameter. The dodecameric cylinder walls have a 30 Å thickness, creating an internal 40 Å diameter pore. The inner surface of the dodecamer is lined with the active sites of the 12 Pdx1 monomers.

For spatial arrangement information of the *Pv*Pdx1 molecules in the crystalline state, the arrangement of the molecules in the unit cell was determined by UCSF Chimera [102]. It shows two molecules, with a distance of approx. 20 Å in the unit cell, as shown in **figure 35**. Further, when the symmetry mates were generated by the visualization program Pymol [172], close packing of the dodecamer molecules became obvious and visible, explaining the good resolution of 2.0 Å **(Figure 36)**.

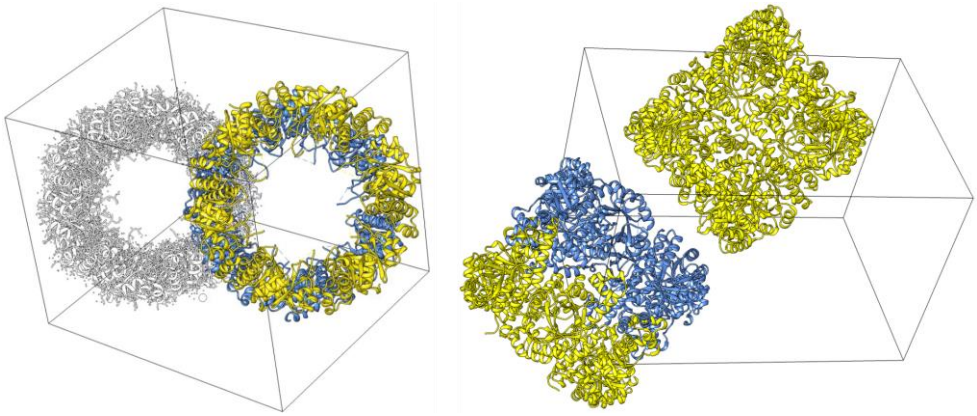


Figure-35. *P.vivax* Pdx1 dodecamer molecules arrangement in the unit cell.
P.vivax Pdx1 molecules arrangement in the unit cell; crystallized in space group P 2_1 occupying 2376650 A^3 unit cell volume, with two space group symmetries packed in a unit cell, shown in front view (left) and side view (right), the figure is prepared by UCSF Chimera [102].

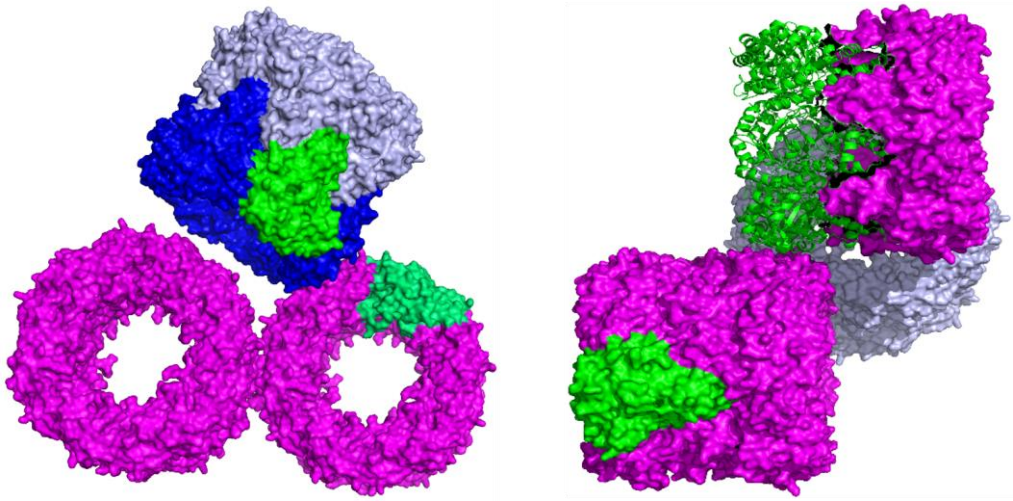


Figure-36. Symmetry mates of the *P.vivax* Pdx1 in the crystal.
For *Pv*Pdx1 one dodecamer is packed in the ASU and two dodecamers in the unit cell. **Left**: front surface view of the three dodecamers mates generated by PyMol [172] represent packing of *Pv*Pdx1 molecules. In a dodecamer, one hexamer is shown light blue and the other in dark blue with a monomer shown in dark green; whereas the other two dodecamers are shown in magenta. *Pv*Pdx1 dodecamers are arranged in a way that two dodecamers are oriented in parallel and the third dodecamer places at an angle, neither parallel and neither perpendicular, providing close packing of the Pdx1 dodecameric assemblies. **Right:** Side view surface model of *Pv*Pdx1 dodecameric spheres composed of 12 monomers, one monomer is shown in green in a dodecamer, in another dodecamer (one hexamer is shown in cartoon representation in green color, whereas the other in the surface form, indicating a very close arrangement of Pdx1 dodecamers from a different angle, in the crystalline form.

4.10.4. Structural analysis of *Pv*Pdx1 monomer $(\beta\alpha)_8$-barrel

The *Pv*Pdx1 structure is composed of 12 chains, with each monomer capable of synthesizing the product PLP. *P.vivax* Px1 monomer sequence consists of 302 amino acid residues. However, the final refined *Pv*Pdx1 structure solved here has a starting 17 N-terminal amino acid residues region found disordered, helix α2′ residues (P-58 to K-64) and last 19 C-terminal residues (M-284 to N-302) were not found in the final electron density. Overall, *Pv*Pdx1 3D structure is revealing a classical $(\beta\alpha)_8$-barrel consisting of 8 β-strands arranged parallel and forming a central β-sheet architecture **(Figure-37a)**. It also comprises of 8 α-helices surrounding the β-barrel with 4 α-helical insertions i.e. α6′, α6′′, α8′, α8′′ and helix αN. The structure possesses some important structural elements, involved in maintaining the assembly, enzyme activity, and interaction with the counterpart (Pdx2) for Pdx complex formation. Firstly, half of the helix αN with residues (L-18 to L-26) is appeared and ordered **(Figure 37a-c)**. This helix is crucial for complex formation after Pdx1 interacts with the Pdx2 subunit; reported in prokaryotes (*B.subtilis*) and Plasmodium [69, 90]. An ordered αN helix was previously only observed in the Pdx1-Pdx2 complexes and disordered in the isolated Pdx1 structures.

Secondly, the helix α2′ insertion between β2 and α2 has a role in enzyme activation [90]. It is not seen in *Pv*Pdx1 like other reported structures [69, 90]. Further, it has been reported that this helix appears after binding the substrate (R5P) and or Pdx2 to the dodecameric Pdx1 [69, 90, 97]. The helix α8′′, which plays a major role in establishing the hexamers, is ordered. Finally, the elongation of the helix α6 and its insertion helices α6′ and α6′′, involved in dodecamer formation are present **(Figure 37a-c)**.

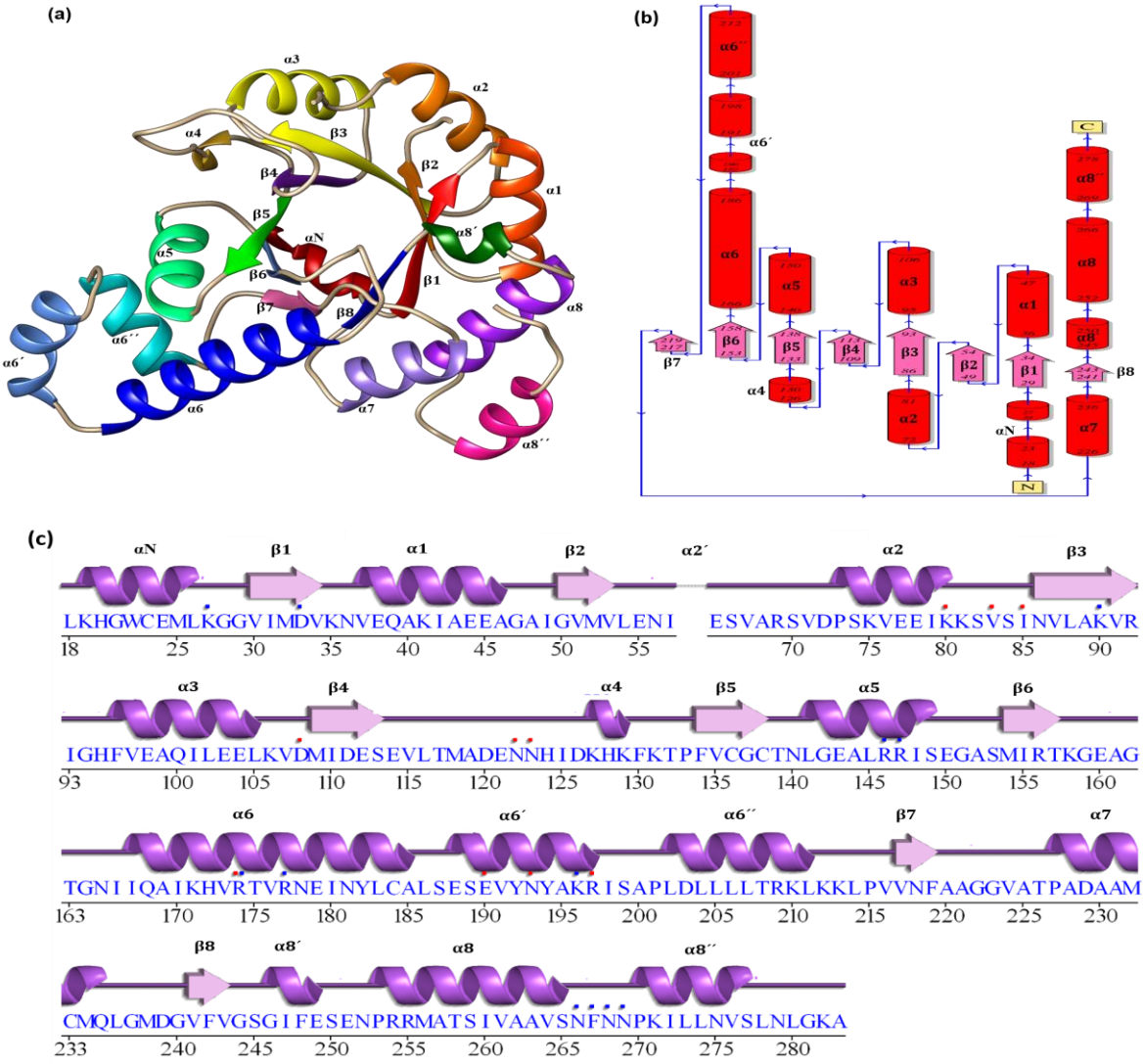


Figure-37. Cartoon representation, topology, and wiring diagram of *Pv*Pdx1 monomer.
(a) *P.vivax* Pdx1 monomer $(\beta\alpha)_8$-barrel with helices and strands shown in different colors and labeled. The structure includes central arranged 8 β-strands surrounded by 13 α-helices, the figure is prepared in UCSF Chimera [102]. **(b)** Schematic diagram illustrating the Pdx1 topology with β-strands (pink arrows) arranged in parallel fashion forming one central β-barrel, and the relative disposition of the α-helices (red cylinders) indicating their 3D arrangement. The N and C terminus along with helices and strands are labeled **(c)** Wiring diagram of *Pv*Pdx1, showing strands (pink arrows), helical elements (purple springs), red dots are the amino acid residues interacting with ligands and blue dots are the metal ions interacting residues. The topology figure and wiring diagram are generated via the online tool PDBsum [104].

Table-16: Data-collection and refinement statistics of *P.vivax* Pdx1

X-Rays Source	**Synchrotron, P11 beamline PETRA III, DESY**
Detector	**Pilatus 6M**
Wavelength (Å)	1.0
Resolution range	48.15 - 2.0 (2.07 - 2.0)
Space group	P 2_1 (4)
Unit cell parameters	
a, b, c (Å)	119.96 167.50 120.20
α, β, ɣ (°)	90 98.46 90
Total reflections	1218276 (118020)
Unique reflections	315045 (31344)
Multiplicity	3.9 (3.8)
Completeness (%)	99.82 (99.76)
Mean I/sigma(I)	11.27 (1.24)
Wilson B-factor	34.21
R-merge	0.0775 (0.8591)
R-meas	0.08987 (0.9999)
$CC_{1/2}$	0.998 (0.633)
Mol. weight in asymmetric unit (kDa)	408.0
Refinement	
Reflections used in the refinement	314940 (31329)
Reflections used for R-free	15749 (1566)
R_{work}	0.1967 (0.3204)
R_{free}	0.2216 (0.3428)
Number of non-hydrogen atoms	24926
Macromolecules	23161
Ligands	205
Solvent	1560
Protein residues	3054
Bond lengths (Å) and angles (°)	0.003 and 0.63
favored (%)	98.07
allowed (%)	1.73
outliers (%)	0.20
Rotamer outliers (%)	0.12
Clash score	2.11
Average B-factor (Å²)	43.09
Macromolecules	42.70
Ligands	64.80
Solvent	46.09

4.10.5. **Interfaces and hexamer formation in *P.vivax* Pdx1**

In the Pdx1 hexamer, one monomer at the interface interacts with two other monomers, e.g. monomer of chain-A has an interface with chain-D and chain-E. The chains A-D have interactions **(Figure 38a,b)** similar to chains A-E, with details provided in table 17. This type of supportive interaction extends between the protomers and result in a stable enzymatically active hexameric structure. The main interacting structural

elements between the protomers at the hexamer interface are α2, α3, α6, α7, the loop region of β3-α3, loop β4-α4, loop β6-α6, loop α2′-α2, and the C-terminus residues. Residues present and involved in interactions between A-D form a total of 9 hydrogen bonds and 4 salt bridges. The residues Arg69, Asp72, Ser74, His95, Arg92, and Asp120 of chain-D with Gly164, Leu280, Asn279, Asp229, and Arg174 of chain-A form single hydrogen bonds. Whereas Gln100 of chain-D is involved in two hydrogen bonds with Gln235 of chain-A. Further, there is a network of salt bridges formed by Arg174 and Asp229 of chain-A and Asp120, Arg69, Arg92, and His95 [175] of chain-D **(Figure 38c)**. The detailed hydrogen bonding and salt bridge interactions of residues from chain-A with chain-D and E protomers are presented in table 17.

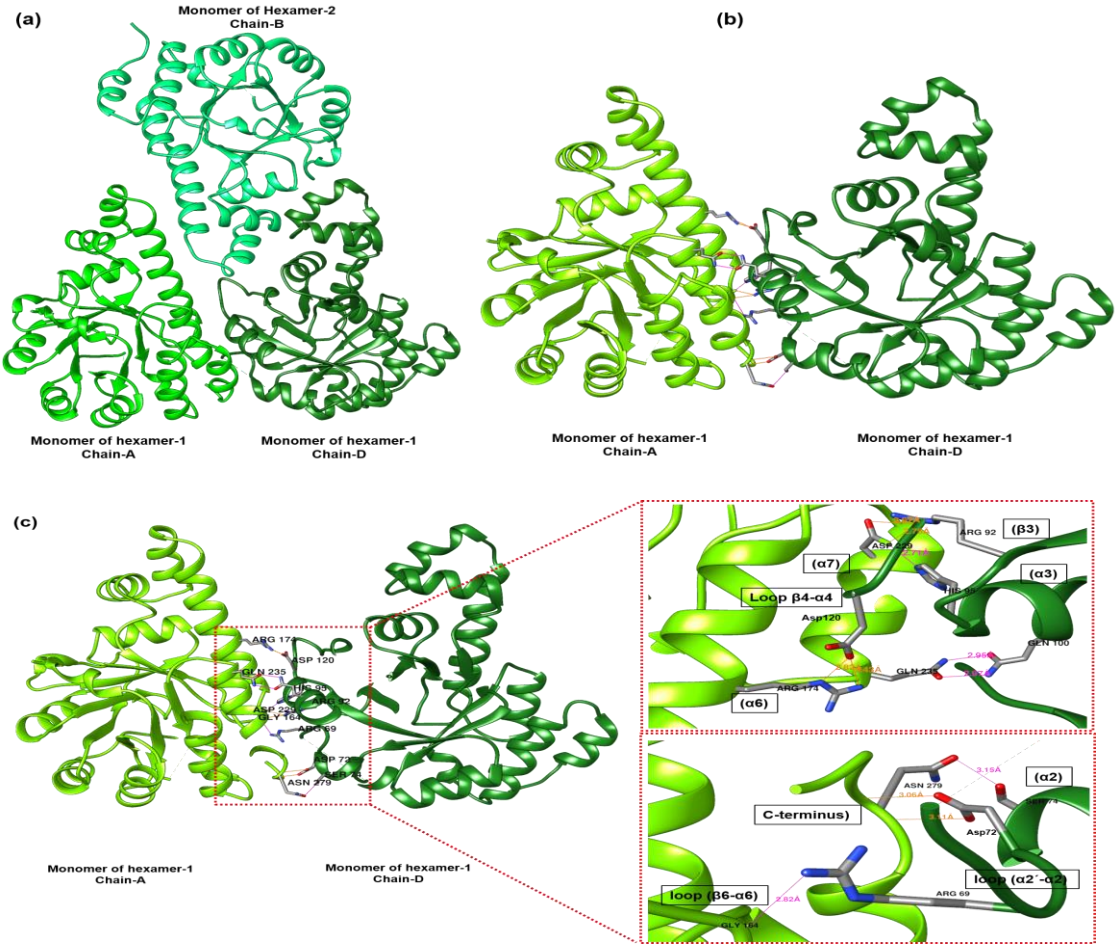


Figure-38. Bonding between protomers involved in *Pv*Pdx1 hexamer formation.
(a,b) *Pv*Pdx1 monomers of chain-A and chain-D in one hexamer and chain B of another hexamer involved in the extensive network of intermolecular hydrogen bonds, salt bridges, and Van-der Waals interactions. **(c) left;** interface of chain-A and chain-D involved in hydrogen bonds, **right;** intermolecular bonding network with the structural elements and position of the interacting residues involved in these interactions is also indicated. The structure figure is prepared by UCSF Chimera [102].

4.10.6. Dodecamer formation in *P.vivax* Pdx1

Dodecameric Pdx1 observed in the asymmetric unit is formed by two enzymatically active Pdx1 hexamers. The C-terminus which is aiding in dodecamer formation is reported disordered in Pdx1, lacking 14–28 amino acid residues [69, 87]. Despite the disordered form in the crystalline state, the last 21 amino acid residues are essential for enzyme activity and become progressively ordered into a loop upon binding a substrate (R5P) and/or Pdx2 [69, 92]. Upon dodecamer formation, a protomer from one hexamer is at the interface to two monomer chains of the second hexamer e.g. chain-A from hexamer-1 forms an interface with chain-B **(Figure 38a, 39)**, also similar to chain-F of hexamer-2 **(Figure 40)**. In other words, one monomer has 4 interfacing chains in a dodecamer, i.e. two interfacing chains in the parent hexamer and the other two interface monomers in the counter hexamer, which result in a dodecamer formation.

The main secondary structure elements involved in the dodecamer formation are helices α6, α6′, and a 3_{10} helix (α4). For interactions between chain-A and chain-B, helices α6 and α6′ are arranged in a parallel manner to expose a maximum surface at the interface. The residues involved in interactions between the two helices are Asn178, Asn181, Glu190 from one chain and Asn178, Ala185, and Glu190 from another chain are connected by hydrogen bonds **(Figure 39)**. The interface additionally includes salt bridging residues Arg174, Arg177, and Glu190 present at helix α6 (Table 17). Further, the two monomers are also involved in 44 Van-der Waals contacts.

On its other side, chain-A has an interface to chain-F and has structural elements with amino acids involved in inter-chain bridging through their main and side chains. Within structural elements, α4 and α6′ residue Ser199 is involved in H-bonding, whereas Asp126, Glu188, Lys129 and Tyr192 are forming salt bridges. Both types of interactions with residues numbers are given in **figure 40** and presented in detail in table 17. Moreover, both chains are also interacting through 46 Van-der Waals contacts.

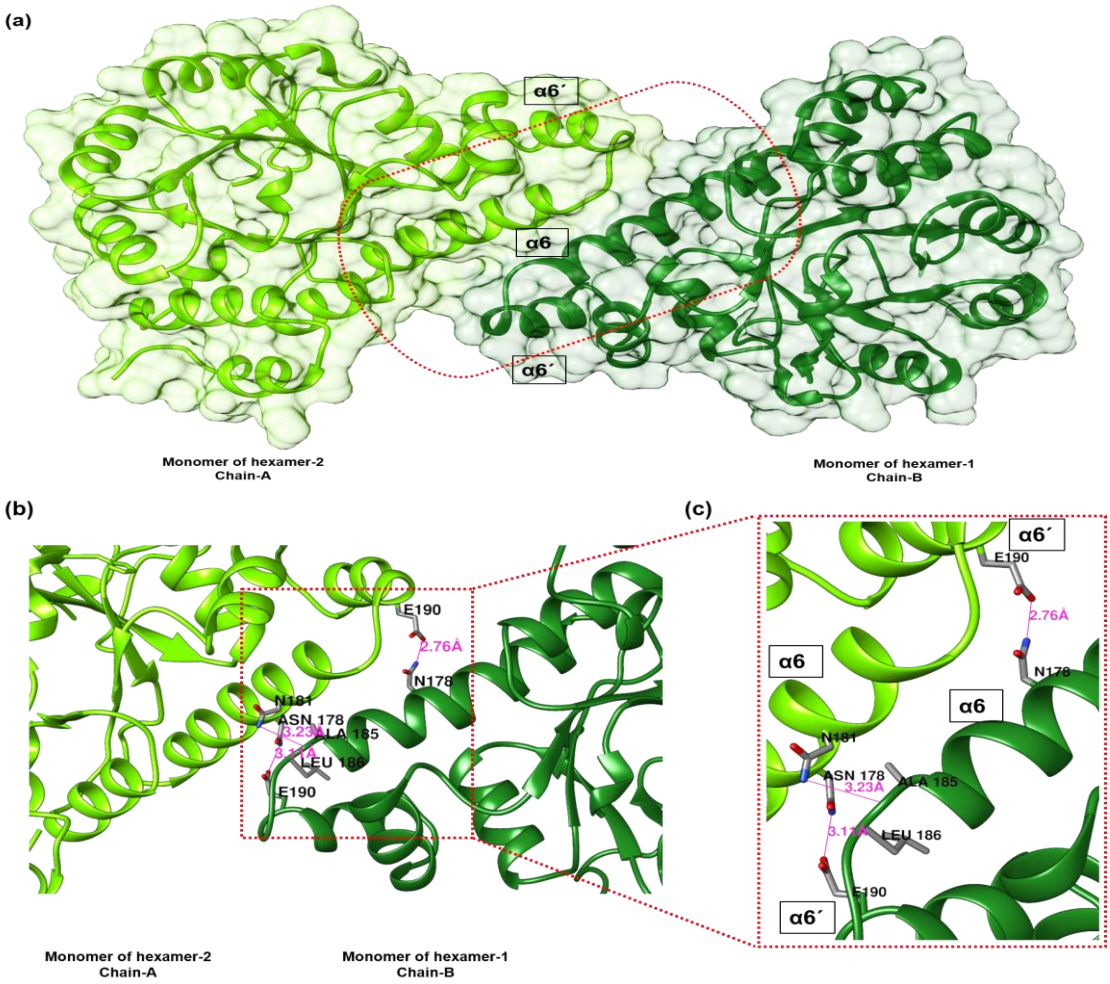
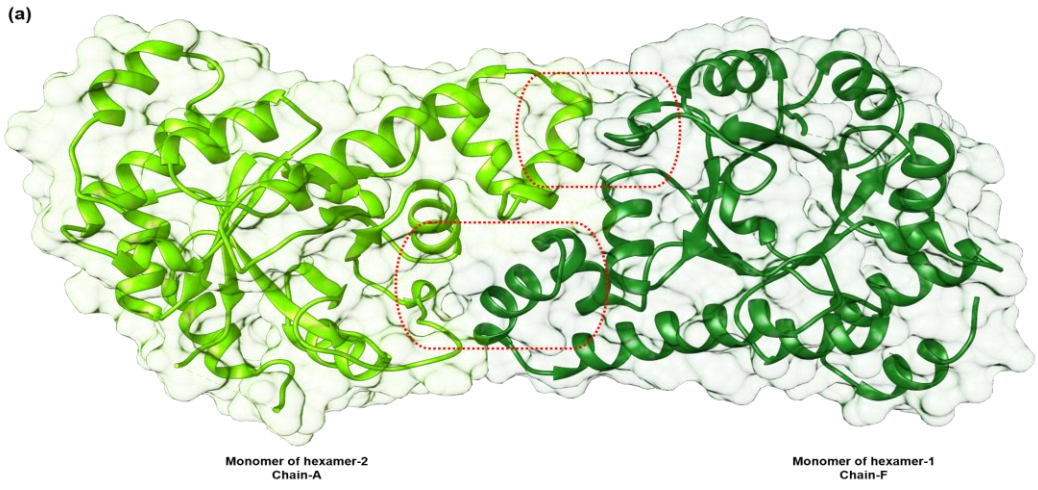


Figure-39. Interactions between *Pv*Pdx1 monomers involved in dodecamer formation.
(a) *Pv*Pdx1 monomer of chain-A from one hexamer and chain-B from another, involved in hydrogen bonding, salt bridges, and Van-der Waals interactions through structural elements α6 and α6´. **(b)** Interface regions and residues of inter-hexamer chain-A and chain-B involved in hydrogen bonding are displayed; **(c)** detailed and zoom-in of residues from helices α6 and α6´ involved in H-bonds with counterpart helices, the figure is prepared by UCSF Chimera [102].



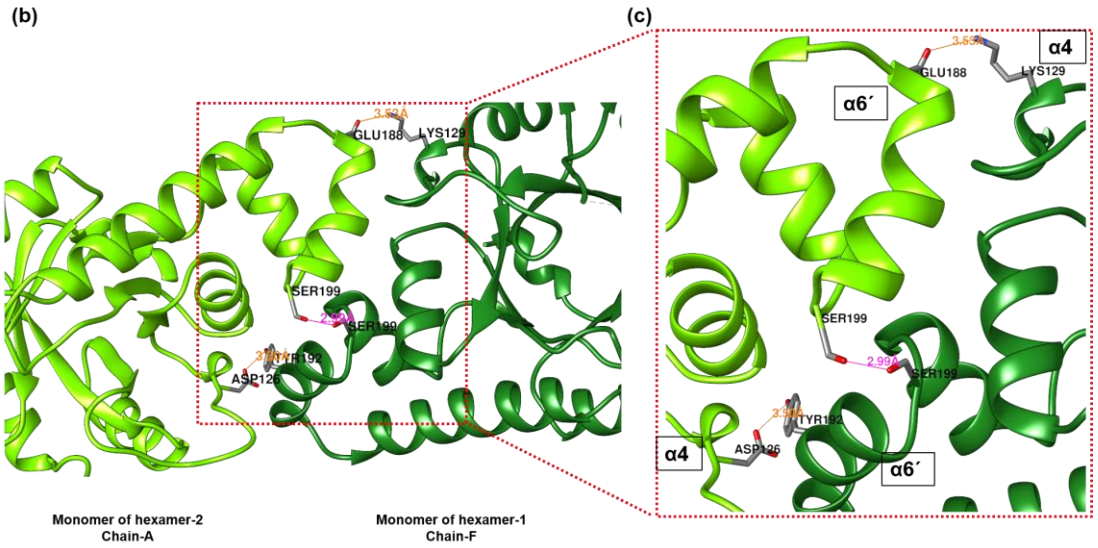


Figure-40. The interface of chain-A and F protomers involved in Pdx1 dodecamer formation. **(a)** *Pv*Pdx1 protomers A and F involved in hydrogen bonding and salt bridges. **(b)** Interface residues of chain-A and chain-F involved in hydrogen and salt bridges; **(c)** residues from helices α4 and α6′ involved in H-bonds and salt bridges are displayed, the figure is prepared by UCSF Chimera [102].

In *Pv*Pdx1 most of the residues are surface exposed and in general, a total of 78 amino acid residues from one monomer (e.g. chain-A) are present at the interface with neighbor protomers (Table 18). The interface residues count approximately 30% of the total of 259 residues in the *Pv*Pdx1 structure. The number of interfaces interacting residues and interface score also known as Complex Formation Significance Score (CSS, the scale of 0.00 to 1.00, with 1.00 indicating a critical role in complex formation) [176], indicates that chains A-D and A-E allow the maximum number of interactions and have a high CSS value. These chains are located within a hexamer, indicating a multimeric state hexamer being more stable than the dodecamer [177] (Table 18). Data for Table 18 were obtained from the online servers PdbSum [104] and PISA [176].

Table-17: Interactions at the interfaces of monomers in *P.vivax* Pdx1

No	Chains	Salt bridge residues	H-bond residues	No of interactions
Hexamer chains				
1	A←→D	Arg174-Asp120	Arg174-Asp120	1
2	A←→D	Arg174-Arg69	Gly164-Arg69	1
3	A←→D	Asp229-His95	Asp229-His95	1
4	A←→D	Asp229-Arg92	Asp229-Arg92	1
5	A←→D	-	Gln235-Gln100	2
6	A←→D	-	Leu280-Asp72	1
7	A←→D	-	Asn279-Asp72	1
8	A←→D	-	Asn279-Ser74	1
9	A←→E	Glu121-Lys171	Arg69-Gly164	1
10	A←→E	Asp120-Arg174	Asp120-Arg174	1
11	A←→E	Arg92-Asp229	Arg92-Asp229	1
12	A←→E	His95-Asp229	His95-Asp229	1
13	A←→E	-	Gln100-Gln235	2
14	A←→E	-	Asp72-Leu280	1
15	A←→E	-	Asp72-Asn279	1
16	A←→E	-	Ser74-Asn279	2
Dodecamer chains				
17	A←→B	Glu190-Arg174	Glu190-Asn178	1
18	A←→B	Arg174-Glu190	Asn178-Glu190	1
19	A←→B	Arg177-Glu190	Asn181-Ala185	1
20	A←→F	Glu188-Lys129	Ser199-Ser199	2
21	A←→F	Asp126-Tyr192	-	-

Table-18: Summary of chain-A monomer interface and solvent accessible data for *P.vivax* Pdx1

Chain-A	Total number of residues	Surface exposed residues	Solvent accessible area ($Å^2$)	Solvation for folding (ΔG, kcal/mol)
	259	231	12089	-244.1
Interfaces interactions calculated by **PdbSum server**				
Chain-A with interface chains	**A→B**	**A→D**	**A→E**	**A→F**
	9:7 interface residues	24:20 interface residues	20:21 interface residues	15:15 interface residues
Interface bonds	3-SB, 3-HB	4-SB, 9-HB	4-SB, 10-HB	1-SB, 2-HB
Interfaces interactions calculated by **PISA server**				
Chain-A with interface chains	**A→B**	**A→D**	**A→E**	**A→F**
	11:13 interface residues	29:25 interface residues	25:26 interface residues	23:24 interface residues
Interface bonds	5-SB, 7-HB	7-SB, 10-HB	7-SB, 10-HB	1-SB, 5-HB
Interface score	0.029	**0.397**	**0.397**	0.143

4.11. *P.vivax* Pdx complex crystallization

P.vivax Pdx complex was purified after mixing SEC pure Pdx1 and Pdx2 proteins in a 1:1 molar ratio at 0.5 mM concentration in the presence of 10 mM L-glutamine. The mixture was then allowed to pass through a Superose 6-Increase 10/300 (GE Healthcare) column with Pdx complex elution. After purification, biophysical and bioanalytical characterization, the *P.vivax* Pdx complex was crystallized at a concentration of 25 mg mL^{-1} and 30 mg mL^{-1} by hanging drop vapor diffusion method. The obtained crystals are shown in **figure 41**. Pdx complex crystallizes in 0.1 M Tris (pH=8.0), 1.0 M LiCl, 9-12% PEG-6000 crystallization buffer. The crystals were cryofreezed in liquid nitrogen utilizing the crystallization buffer with an additional 20% glycerol for cryo-protection. After extensive cryoprotection optimization, Pdx complex crystals were used for X-ray data collection at beamlines P11 and P13. However, most of them did not diffract and only a few diffracted to a very low resolution of approximately 10 Å.

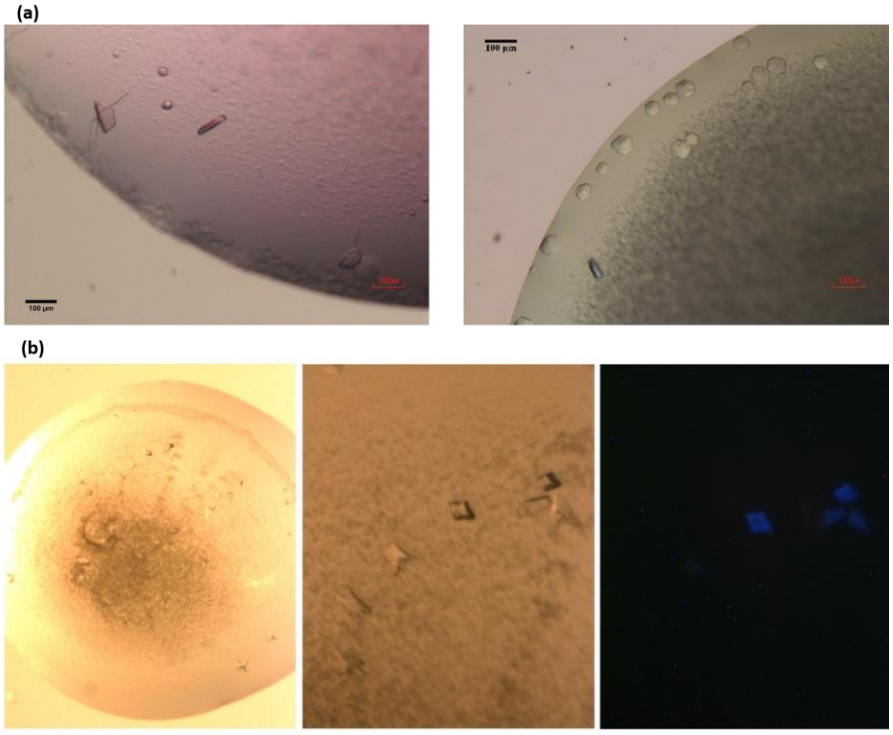


Figure-41. *P.vivax* Pdx complex crystals obtained and confirmed by exposure with UV light. **(a)** *Pv*Pdx complex crystals before freezing with a cryoprotectant, grown in one dimension to approx. 60 µm. **(b)** Pdx complex crystals, UV confirmation applying the Crystal Score device (Diversified Scientific Inc, USA).

4.12. *S.aureus* Pdx1 crystallization

After extensive high-throughput crystallization screening Pdx1 from *S.aureus* crystallizes by sitting drop crystallization method under conditions of the following crystallization screens. JCSG screen with optimal crystals in conditions B12 and G6, PACT screen in G9, H9, H11, H12, and Morpheus conditions C1, C2, C9, and C12 applying the honeybee robot. The crystals appeared at a concentration of 12 mg mL^{-1}. Crystals were cryo-cooled in liquid nitrogen utilizing crystallization buffer and an additional 30% of glycerol for *Sa*Pdx1 crystals for cryo-protection. The flash-cooled crystals, shown in **figure 42a,b,** were used for X-ray diffraction and data collection. Diffraction data were collected at beamline P11, PETRA-III (DESY, Hamburg). Details of the collected and processed data for *Sa*Pdx1 are summarized and presented in table 19.

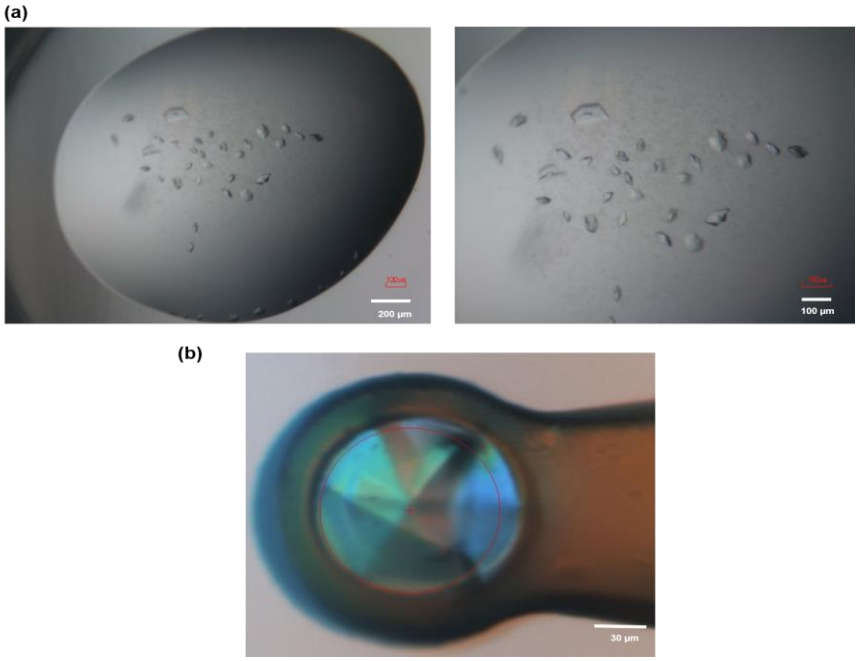


Figure-42. *S.aureus* Pdx1 crystals grown and optimized.
(a) *Sa*Pdx1 crystals with dimensions of 120 x 90 µm before freezing with a cryoprotectant **(b)** *Sa*Pdx1 crystal diffracted and processed at a resolution of 2.83 Å, a snapshot was taken right before the data collection at beamline P11 PETRA-III, DESY.

4.12.1. *S.aureus* Pdx1 octamer in ASU

The *Sa*Pdx1 diffraction data were processed by using the software XDSAPP [120] to obtain the reflection (mtz) file. Applying the Matthews probability calculator [174], molecular mass information, and asymmetric unit (ASU) information were obtained. *Sa*Pdx1 crystallizes in the rhombohedral/trigonal space group (R 3 2), and its packing

remains such that the monomers do not pack into one complete hexamer or dodecamer in the asymmetric unit. The Matthews parameter suggested a hexamer with 70% solvent content in the asymmetric unit (dodecamer of molecular weight 384 kDa was suggested with 40% solvent content). However, the final 3D structure of *Sa*Pdx1 was solved with an octamer per asymmetric unit occupying 14302472 A^3 volume **(Figure 43)**. *Sa*Pdx1 octamer packed in the asymmetric unit with a molecular weight of 256 kDa with a corresponding solvent content of 60% in the ASU. The Pdx1 octamer packed in ASU electron density is shown in **figure 43a,b**, with potential dodecamers identified from the electron density distribution in **figure 43b,** right.

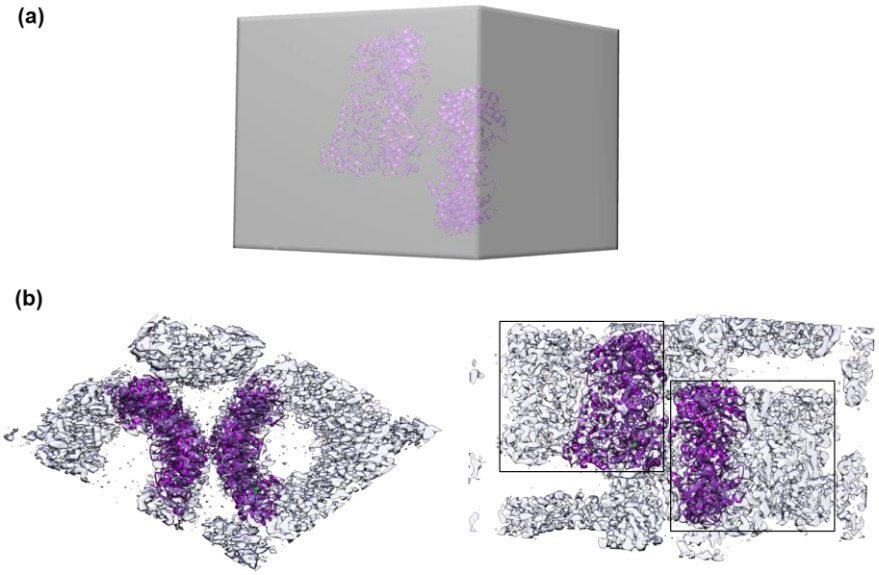


Figure-43. *S.aureus* Pdx1 crystalized and observed as octamer in the asymmetric unit.
(a) Pdx1 molecules crystallized in ASU with 4mers belong to one hexamer/dodecamer unit and other 4mers to the other oligomeric unit **(b) left:** Pdx1 octamer (in magenta color) top view with electron density light grey in the asymmetric unit. **Right:** Pdx1 octamer (magenta) side view with an electron density of ASU occupying a volume of 794581 A^3, suggesting dodecamer formation in the crystalline form with two dodecamers encircled. The suggested dodecamers are observed well separated from each other. The figure is prepared by UCSF Chimera [102].

4.12.2. *Sa*Pdx1 oligomeric structure and arrangement

Diffraction data for *Sa*Pdx1 were processed using the XDS suite [120]. The 3D structure was determined by molecular replacement, using *Bacillus subtilis* Pdx1 (PDB code: 2NV1) [90] with 80% sequence identity as a reference model, applying the program Phaser-MR [158]. The structure model was refined by the program called Phenix-refine [159] and model building was carried out using the program Coot [160] with resulting R_{work} of 20.59 %, R_{free} of 25.28 %, and a total of 96 solvent molecules,

as given in table 19. The final refined *Sa*Pdx1 octamer packed in the unit cell is shown in **figure 44**. The unit cell contains 18 octameric asymmetric units **(Figure 44)**. In other words, 12 *Sa*Pdx1 dodecameric symmetries are packed in a unit cell, accumulating a total mass of 4608 kDa. One ASU occupies a volume of 794581 A^3 and 18 symmetries make a unit cell volume of 18 x 794581 A^3 equal to 14302472 A^3.

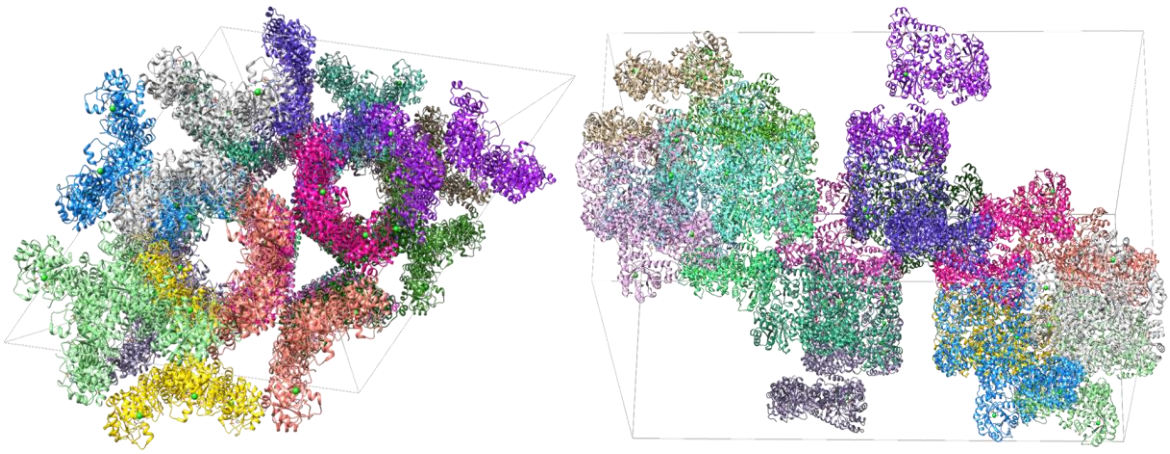


Figure-44. *S.aureus* Pdx1 molecules complicated arrangement in the unit cell.
S.aureus Pdx1 molecules arrangement in a unit cell; *Sa*Pdx1 is crystallized in space group R 3 2: H with a unit cell volume of 14302472 A^3 containing 18 space group symmetries, each with a volume of 794581 A^3 are packed in a unit cell, shown in the front view (left) and side view (right) with potential dodecamer molecular assemblies. Figure is prepared by UCSF Chimera [102].

The ASU electron density **in figure 43** and unit cell packing of *Sa*Pdx1 **(Figure 44)** are suggesting a potential dodecameric form of *Sa*Pdx1 in the crystalline state. Therefore, symmetry mates were generated with the molecular visualization program PyMol [172], and a dodecamer formation in *Sa*Pdx1 was observed, as shown in **figure 45**. The resulting dodecamers are apart with a distance of approximately 26 Å in a unit cell with an arrangement having more space in between them **(Figure 45)**. This dodecameric assumption requires further detailed structural investigation of the crystal structure to identify potential differences from *Pv*Pdx1 and other reported Pdx1 homologs. As it is showing a hexamer state in solution, whereas a different oligomeric state in the crystal packing.

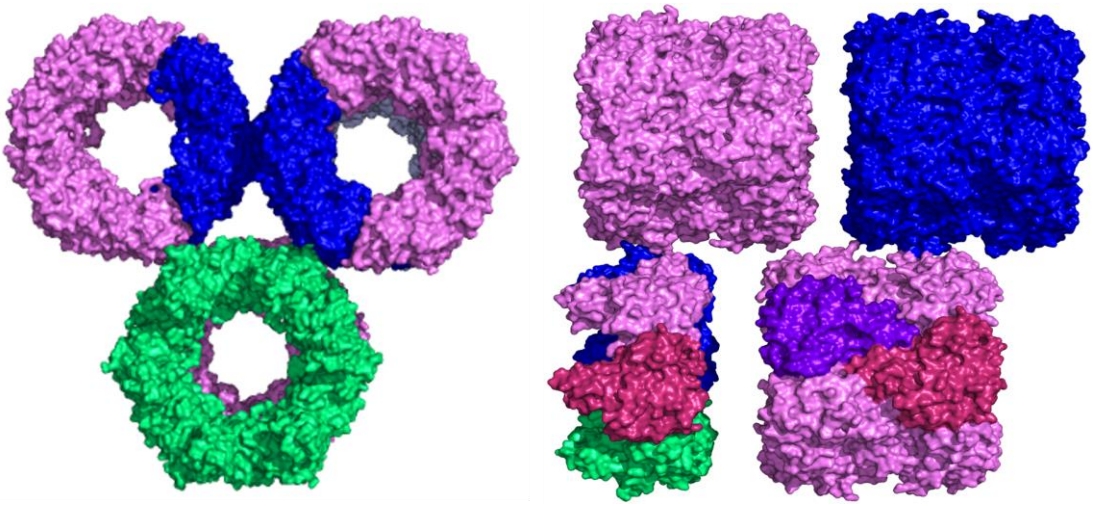


Figure-45. Symmetry mates of *S.aureus* Pdx1 indicating dodecamers in the unit cell.
The **(left)** surfaces generated from symmetry mates indicate front view packing of *Sa*Pdx1 dodecameric molecules. *Sa*Pdx1 one octameric ASU packing (blue) in dodecamers, where half of ASU is part of one and half of another dodecamer. **Right** surface view provides the side view surface representation of *Sa*Pdx1 dodecamers arranged in parallel fashion, showing open packing of Pdx1 dodecamer molecules with the inter-dodecamer distance of approx. 26 Å. *Sa*Pdx1 dodecameric spheres are composed of 12 monomers (shown in the bottom right) with two monomers shown in different colors, indicating protomers arrangement in a hexamer and dodecamer. Symmetries mates are generated by PyMol [172].

4.12.3. Important structural elements in the *Sa*Pdx1 monomer (βα)$_8$-barrel

The *Sa*Pdx1 crystal structure is composed of eight chains (octamer) with four chains are part of one and the other four are part of another dodecamer. *S.aureus* Px1 with UniProt sequence (Uniprot id: P60798) consists of 295 amino acid residues, whereas in the refined *Sa*Pdx1 structure 273 residues (I-5-K-277) appeared and solved **(Figure 46a-c)**. The structure starting four N-terminal residues and 18 C-terminal residues regions were not found in the final electron density. The complete α-helix (αN) with residues (5-18) is well folded and ordered **(Figure 46a-c)**. An ordered αN helix was previously only observed in Pdx1-Pdx2 complexes, but not found in the isolated Pdx1 subunits. Second, the ordered helix α2′ located between β2 and α2, involved in the enzyme activation, appeared and folded well in *Sa*Pdx1. Whereas α2′ is disordered in *P.vivax* Pdx1 **(Figure 37a)** and also in other plasmodial species [69], as well as in the reported bacterial homologs e.g. *Bs*Pdx1 [90]. It is reported that this loop appears after binding of the substrate (R5P) and or Pdx2 to Pdx1 [69, 90, 97].

Overall, *Sa*Pdx1 monomers consist of eight β-strands arranged parallel and forming a central β-sheet architecture. It also comprises of eight α-helices surrounding the β-

barrel with five α-helical insertions i.e. α6′, α6′′, α8′, α8′′, a novel α2′ and helix αN, thus presenting a classical $(\beta\alpha)_8$-barrel architecture **(Figure 46a)**.

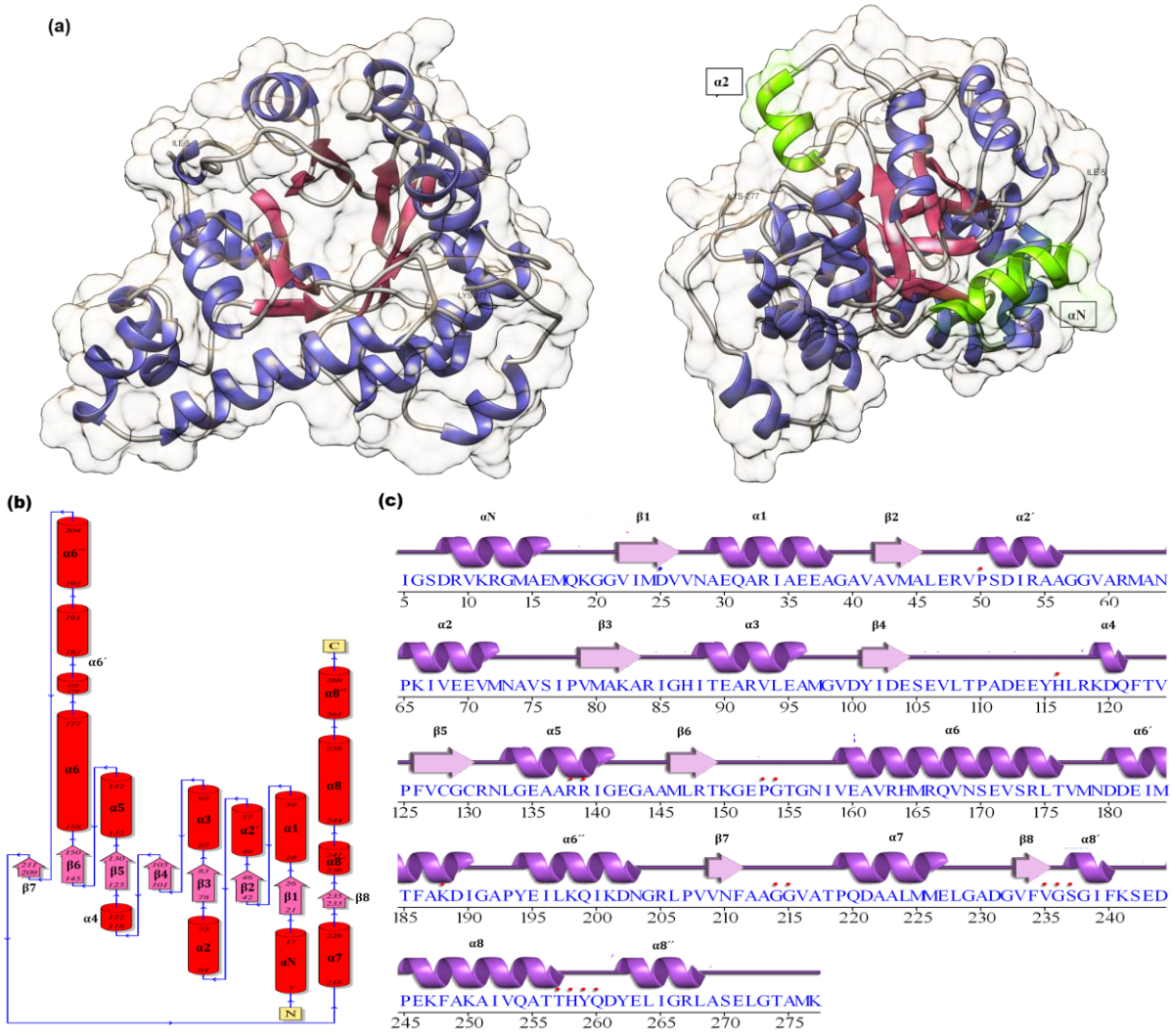


Figure-46. Cartoon representation, topology, and wiring diagram of *S.aureus* Pdx1.
(a) *Sa*Pdx1 monomer $(\beta\alpha)_8$-barrel with secondary structure elements are shown in different colors (helices are shown in blue, β-strands in purple, and loops represented in light green color), N and C-terminus are labeled. The structure includes 8 β-strands arranged in the center, surrounded by 14 α-helices. The structure figure is prepared by UCSF chimera [102]. Right: helix α2′ and αN spatial positions in 3 dimensions are shown in green. **(b)** Schematic diagram illustrating the Pdx1 topology indicating the β-strands (pink arrows) arranged in parallel forming one central β-barrel, and the α-helices (red cylinders) arrangement, indicating their 3D arrangement in the structure. The N and C terminus with secondary structural elements (helices and strands) are labeled, and residue numbers for the beginning and end of each structural element are mentioned. **(c)** Wiring diagram of *Sa*Pdx1 secondary structure including strands (pink arrows), helical elements (purple springs), red dots are the amino acid residues interacting with ligands and blue dots are the metal ions interacting residues. The topology figure and wiring diagram are prepared with PdbSum [104].

Table-19: Data-collection and refinement statistics of *S.aureus* Pdx1

X-Ray Source	Synchrotron, beamline P11, PETRA III, DESY
Detector	Pilatus 6M
Wavelength (Å)	1.0
Resolution range (Å)	49.8-2.82 (2.92-2.82)
Space group	R 3 2: H
Unit cell parameters	
a, b, c (Å)	192.41 192.41 448.19
α, β, γ (°)	90 90 120
Total number of reflections	782916 (73691)
Unique reflections	76303 (7444)
Multiplicity	10.3 (9.9)
Completeness (%)	99.59 (98.36)
Mean I/sigma (I)	16.04 (3.54)
Wilson B-factor	56.98
R_{merge}	0.1081 (0.5976)
R_{meas}	0.1138 (0.6303)
$CC_{1/2}$	0.999 (0.954)
Mol. weight in the asymmetric unit (kDa)	256.0
Refinement	
Reflections used for refinement	76169 (7428)
Reflections used for R_{free}	3811 (372)
R_{work}	0.2059 (0.3069)
R_{free}	0.2528 (0.3660)
CC_{work}	0.954 (0.905)
CC_{free}	0.942 (0.806)
Number of non-hydrogen atoms	16449
Macromolecules	16255
Ligands	98
Solvent Molecules	96
Protein residues	2165
RMS (bonds) (Å)	0.004
RMS (angles) (°)	0.64
Ramachandran favored (%)	96.18
Ramachandran allowed (%)	3.08
Ramachandran outliers (%)	0.75
Rotamer outliers (%)	0.00
Clashscore	6.27
Average B-factor (Å²)	66.33
Macromolecules	66.36
Ligands	70.13
Solvent	57.73

4.12.4. *S.aureus* Pdx1 bonding network in hexamer

In a *Sa*Pdx1 hexamer, one monomer is at the interface of two other protomers e.g. monomer of chain-A is at the interface of chain-E and chain-F having almost similar interactions with both chains, as shown in **figure 47a,b**, and further details are

provided in table 20. The prime interacting structural elements between the monomers at the hexamer interface are α2´, β3, α3, α6, α7, loop region of β4-α4, loop β6-α6, and of the C-terminus residues. A total of 21:18 residues are at the interface in chain-A and E in a hexamer, involved in a total of 7 hydrogen bonds and 5 salt bridges with 87 Van-der Waals interactions. The residues Asp221, Gly156, Arg163, Arg166, Gln220, and Ala275 of chain-A are involved in single hydrogen bonds with Arg84, His87, Arg61, Glu113, Asp112, and Thr89 of chain-E **(Figure 47b,c)**. The Arg163, Arg166 and Asp221 of chain-A are involved in salt bridges with Glu113, Glu114, Asp112, Arg84 and His87 [175]. The detailed hydrogen bonds and salt bridges between residues of different protomers in a hexamer formation are presented in table 20.

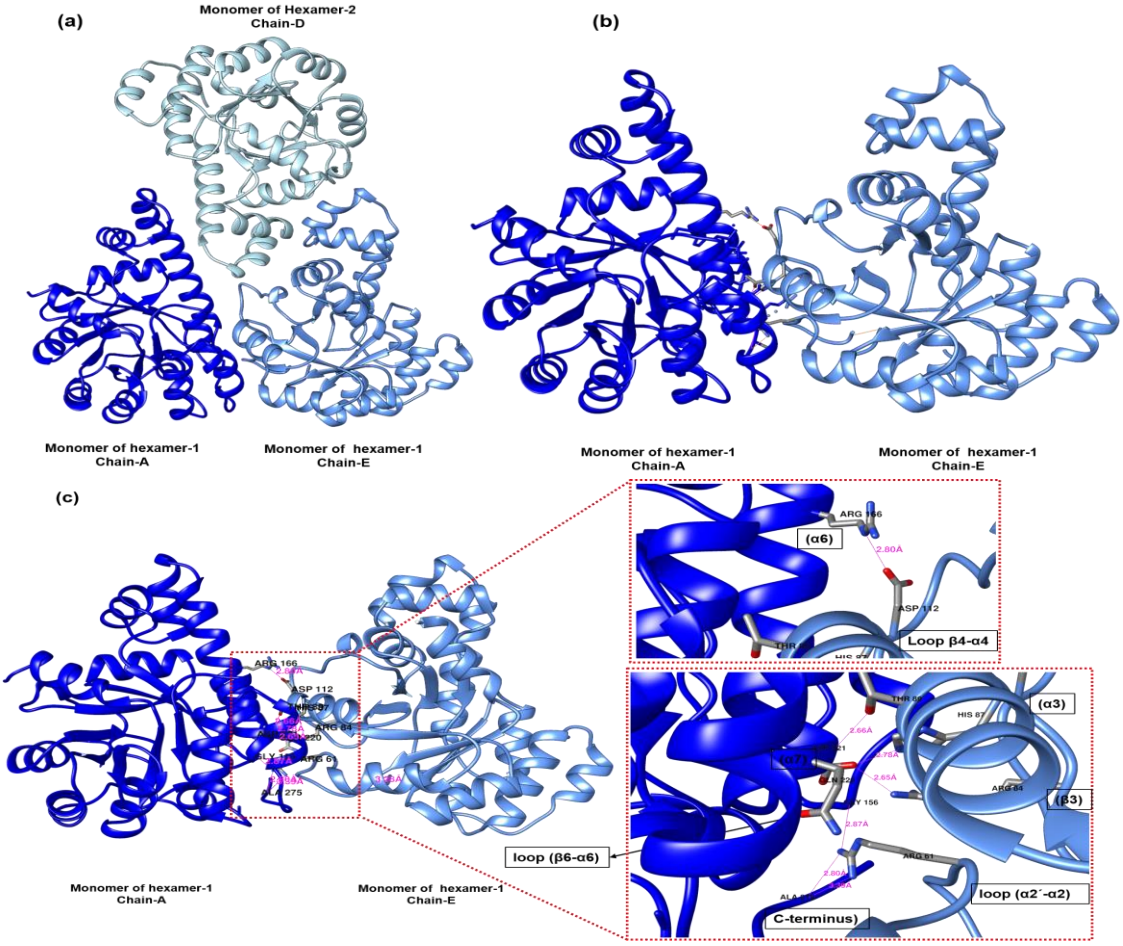


Figure-47. Interactions in *S.aureus* Pdx1 protomers involved in hexamer formation.
(a,b) *Sa*Pdx1 monomers chain-A and chain-E interface region involved in an extensive network of intermolecular hydrogen bonding, salt bridges, and van-der Waals contacts; **(c) left;** interface of chain-A and chain-E involved in hydrogen bonds, **right;** intermolecular bonding network with the structural elements and position of the residues involved in these bonding are also indicated. The figure is prepared by UCSF Chimera [102].

4.12.5. *S.aureus* Pdx1 bonding network in dodecamer

*Sa*Pdx1 exists in a hexamer form in solution but as a dodecamer is also observed in the crystalline form, as shown in **figures 43-45**. Therefore, the solved crystal structure was investigated to identify detailed interchain intermolecular interactions allowing a dodecamer formation. **Figure 48a-c** indicates that chain-A from one hexamer (hexamer-1) is at the interface of chain-D from another hexamer (hexamer-2). The main secondary structure elements involved in the dodecamer formation are helices α6 and α6′ and 3_{10} helix α4, similar to what was observed in the *P.vivax* Pdx1 structure. Helices α6 and α6′ are arranged in a parallel fashion allowing a maximum number of interactions between both the chains. The residues involved in bonding between the two helices are Arg166, Glu182 of chain-A with Arg166, Asp181, and Glu182 of chain-D forming H-bonds as well as salt bridges **(Figure 48b,c** and table 20). Further, the two monomers are also involved in 26 Van-der Waals interactions.

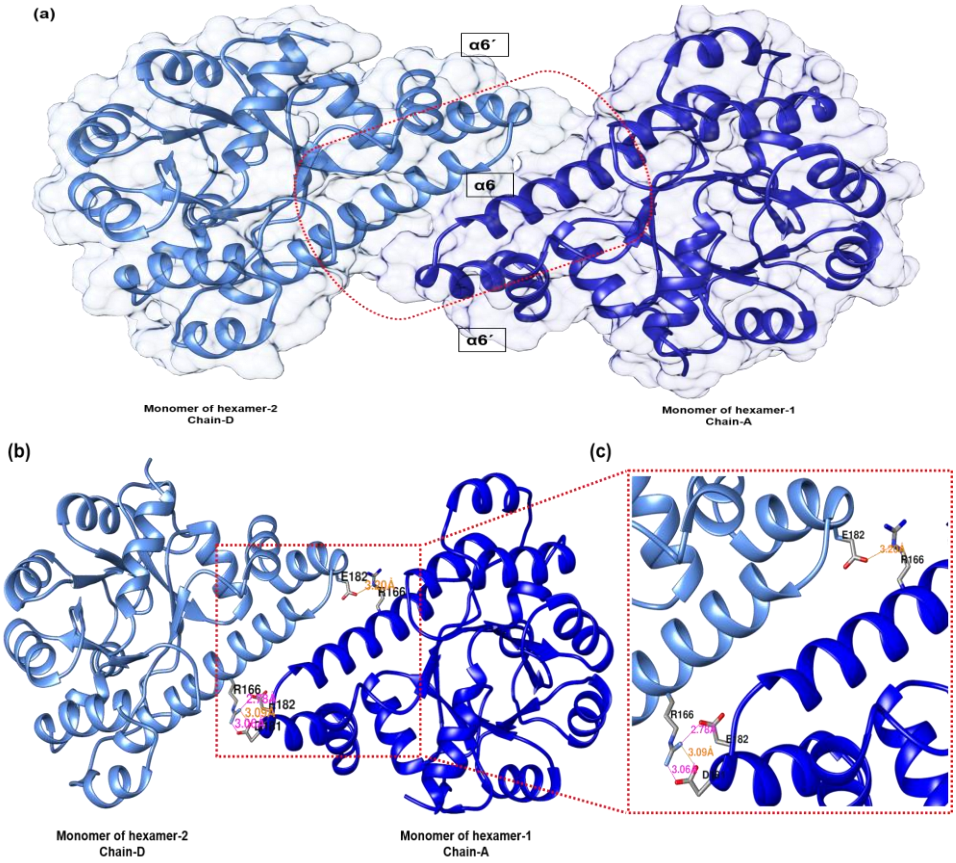


Figure-48. Interactions between *Sa*Pdx1 monomers involved in dodecamer formation
(a) *Sa*Pdx1 monomers of chain-A from one hexamer and chain-D from another are at interfaces with structural elements α6 and α6′ involved in inter-chain bonding **(b)** Interface residues of chain-A and chain-D involved in hydrogen bonds are shown; **(c)** detailed and zoom-in of residues from helices α6 and α6′ involved in H-bonds and salt bridges with counterpart helices. The figure is prepared by UCSF Chimera [102].

Subsequently, the monomer of chain-A was also investigated for interface bonding with chain-B present at its dodecameric interface and a few residues for potential inter-chain bonding were found. A total of 19:17 residues are located at the interface of chain-A and B which are involved in a total of five hydrogen bonds with no salt bridge residues were found, contrary to what was observed for the *P.vivax* Pdx1 dodecamer. The structural elements involved in inter-hexamer interactions are α4, α6′ and α6′′ **(Figure 49a)** with interacting residues Glu114, Arg118, Gln121, Asp180, Glu195 of chain-A with Thr185, Asp180, Arg118, and Glu195 of chain-B involved in H-bonds. Details are presented in **figure 49b-c** and table 20. Moreover, both chains are also connected through 98 Van-der Waals contacts.

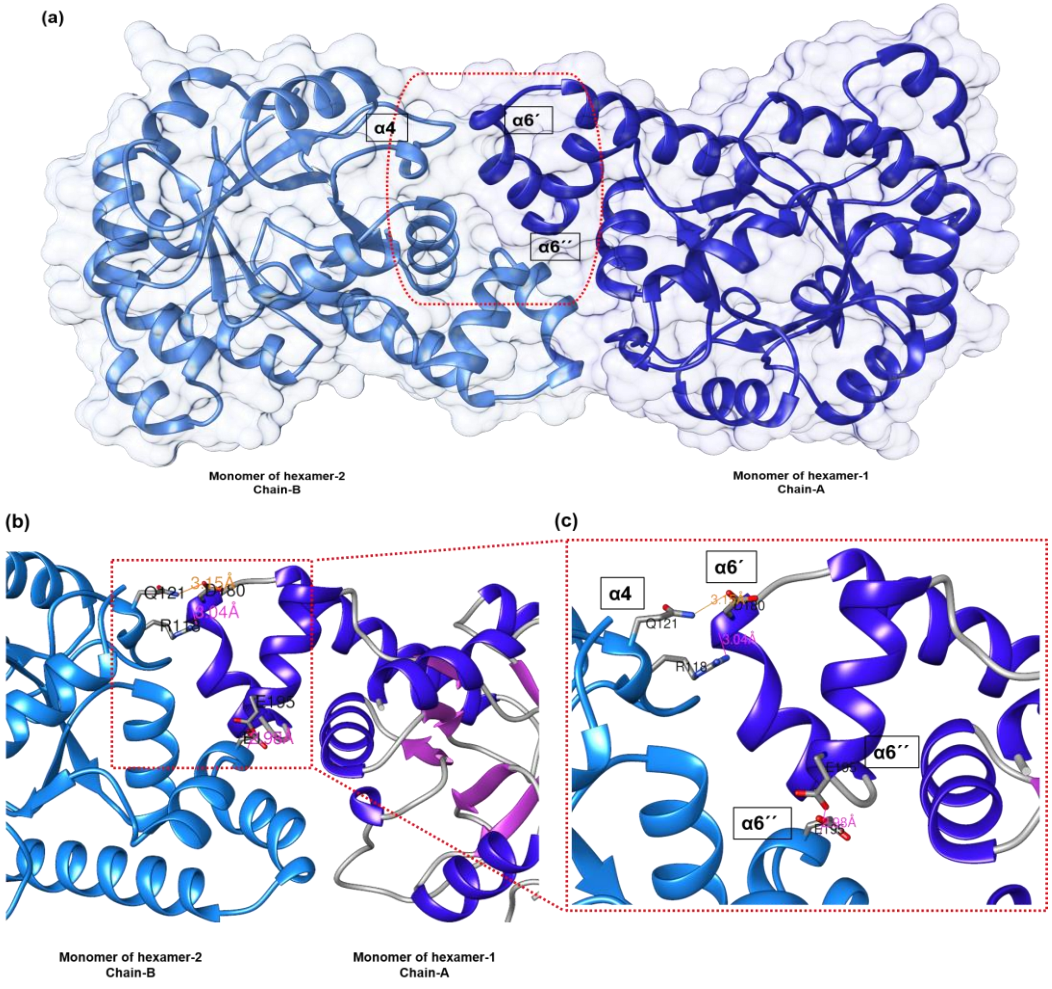


Figure-49. The dodecameric interface of chain A and B monomers.
(a) *Sa*Pdx1 monomers of chain-A and chain-B are involved in hydrogen bonding with mentioned secondary structural elements **(b)** Interface residues of chain-A and chain-B involved in bonding are shown; **(c)** residues from helices α4 and α6′ and α6′′ involved in H-bonds formation are shown. The figure is prepared by UCSF Chimera [102].

In *Sa*Pdx1 a total of 241 residues are surface exposed and in general, a total of 74 amino acid residues from one monomer (e.g. chain-A) are present at the interface of other neighbour monomers (Table 21). These interface residues make 27% of the total 273 residues of the *Sa*Pdx1 monomer. The number of interfaces interacting residues and interface score (CSS) indicates that chains A-E have a maximum CSS value of 0.350 (Table 21). Both these chains are involved in the formation of a hexamer, indicating that the hexamer form is more stable than its dodecamer form. The data shown in table 21 are obtained from online servers PdbSum [104] and PISA [176].

Table-20: Interactions at the interfaces of monomers in *S.aureus* Pdx1

No.	Chains	Salt bridge residues	H-bond residues	No of interactions
Dodecamer chains				
1	A←→B	-	Glu114-Thr185	1
2	A←→B	-	Arg118-Asp180	1
3	A←→B	-	Gln121-Asp180	1
4	A←→B	-	Asp180-Arg118	1
5	A←→B	-	Glu195-Glu195	1
6	A←→D	Arg166-Asp181	Arg166-Asp181	2
7	A←→D	Arg166-Glu182	Arg166-Glu182	1
8	A←→D	Asp181-Arg166	-	1
9	A←→D	Glu182-Arg166	Glu182-Arg166	1
Hexamer chains				
10	A←→E	Asp221-Arg84	Asp221-Arg84	1
11	A←→E	Asp221-His87	Asp221-His87	1
12	A←→E	Arg163-Glu113	Arg163-Glu113	1
13	A←→E	Arg163-Glu114	Gln220-Thr89	1
14	A←→E	Arg166-Asp112	Arg166-Asp112	1
15	A←→E	-	Gly156-Arg61	1
16	A←→E	-	Ala275-Arg61	1
17	A←→F	Arg84-Asp221	Val59-Ala275	1
18	A←→F	-	Arg61-Gly156	1
19	A←→F	-	His87-Asp221	1
20	A←→F	-	Thr89-Gln220	1
21	A←→F	-	Asp112-Arg166	1

Table-21: Summary of chain-A interface and solvent accessible data for *S.aureus* Pdx1

Chain-A	Total number of residues	Surface exposed residues	Solvent accessible area ($Å^2$)	Solvation for folding
	273	241	11858	273
Interfaces interactions calculated by **PdbSum server**				
Chain-A with interface chains	**A→B**	**A→D**	**A→E**	**A→F**
	19:17 interface residues	7:7 interface residues	21:18 interface residues	17:13 interface residues
Interface bonds	5-HB	4-SB, 5-HB	5-SB, 7-HB	1-SB, 5-HB
Interfaces interactions calculated by **PISA server**				
Chain-A with interface chains	**A→B**	**A→D**	**A→E**	**A→F**
	26:26 interface residues	11:12 interface residues	25:31 interface residues	21:23 interface residues
Interface bonds	5-HB	4-SB, 5-HB	5-SB, 7-HB	1-SB, 5-HB
Interface score	0.169	0.035	**0.350**	0.153

4.12.6. Dodecamer formation in S.*aureus* Pdx1

*Sa*Pdx1 is found as a hexamer in solution, indicated by different biophysical and bioanalytical techniques, whereas its crystal form shows a dodecameric state. The helix a6´, an important structural element required for dodecamer formation, is shorter in *S.aureus* Pdx1 from the N-terminal side compared to *Pv*Pdx1, which results in a smaller number of contacts (Table 23). Further, the *Pv*Pdx1 residues Asn178, Asn181 that form inter-hexamer bonds are mutated to Ser170, Ser173 in *Sa*Pdx1 and resulting in a loss of H-bonds. Similarly, *Pv*Pdx1 residue Ser199 of one chain is involved in two hydrogen bonds with Ser199 of the interface chain, is a Gly191 in *Sa*Pdx1 do not form H-bonds to the interface chain. Further, *S.aureus* Pdx1 residues Met184 (*Pv* Tyr192) and Asn169 (*Pv* Arg177) form no salt bridges, whereas in *Pv*Pdx1 the residues at the corresponding positions are involved in interactions (Table 17, 20). The loss in hydrogen bonding and salt bridges could be the reason for the hexameric state of *S.aureus* Pdx1 in solution. Contrary to the hexameric state, *Sa*Pdx1 is a dodecamer in the crystalline state. The reason is the presence of a phosphate ion (PO_4) located at the interface of the dodecameric chains (e.g. A and B) at the P2 site. This generates a hydrogen-bonding network that connects both chains, thus forming a dodecamer

(Figure 50). These ions connect helix α6′ of one chain to helix α5 of another chain. They are potentially originated from the crystallization condition, where they are present in form of phosphate and sulfate salts of different metals.

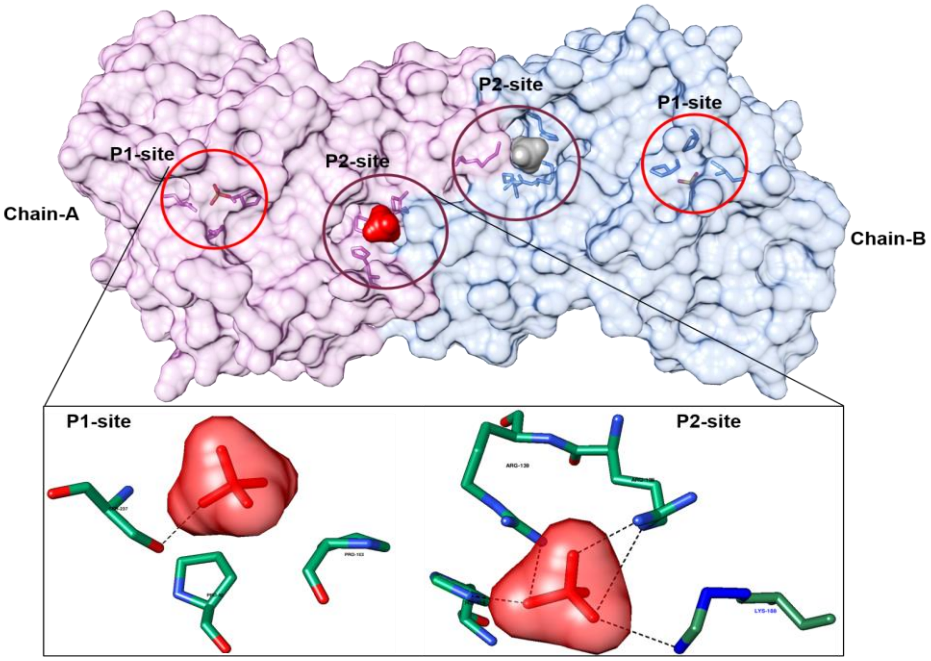


Figure-50. Phosphate ion present at the hexamer interface of *S.aureus* Pdx1 involved in dodecamer formation.
Protomers from chains A and B are connected through PO_4 ions present at the inter-chain hexamer interface at the P2 site, located at a distance of 20 Å from site P1. The PO_4 ions connect Lys188 from helix α6′ of chain-B to the PLP binding site residues Arg138, Arg139 in the helix α5 of chain-A, through hydrogen bonds and form the dodecamer.

4.13. *P.vivax* and *S.aureus* Pdx1 structures comparison

The amino acid sequence of *P.vivax* Pdx1 was aligned to the sequence of *S.aureus* Pdx1 along with their respective closest homologous Pdx1 sequences from *P.berghei* (85% sequence identity), and *B.subtilis* (80% sequence identity), as shown in **figure 51**. Further, both solved structures (*P.vivax* and *S.aureus* Pdx1) have 52% sequence identity to each other, present almost at the borderline to be called homologous structures. The sequence comparison indicates that the domains involved in the dodecamer formation i.e. loop region β4-α4, and helices α4, α6 and α6′ are not strictly conserved and are specie specific **(Figure 51** and Table 22, 23). The results of the comparison also indicate that N-and C-terminals are the most non-conserved regions in the sequences in addition to the helix α8′′, which is involved in protomers interactions within the hexamer. In the *P.vivax* Pdx1 sequence, the elongated N-terminus region was not found in the electron density. This *Pv*Pdx1 elongated

sequence, which is not present in its homologous structures **(Figure 51)**, offers an open question for its structural elucidation and function with the presumptive role in the Pdx complex stabilization.

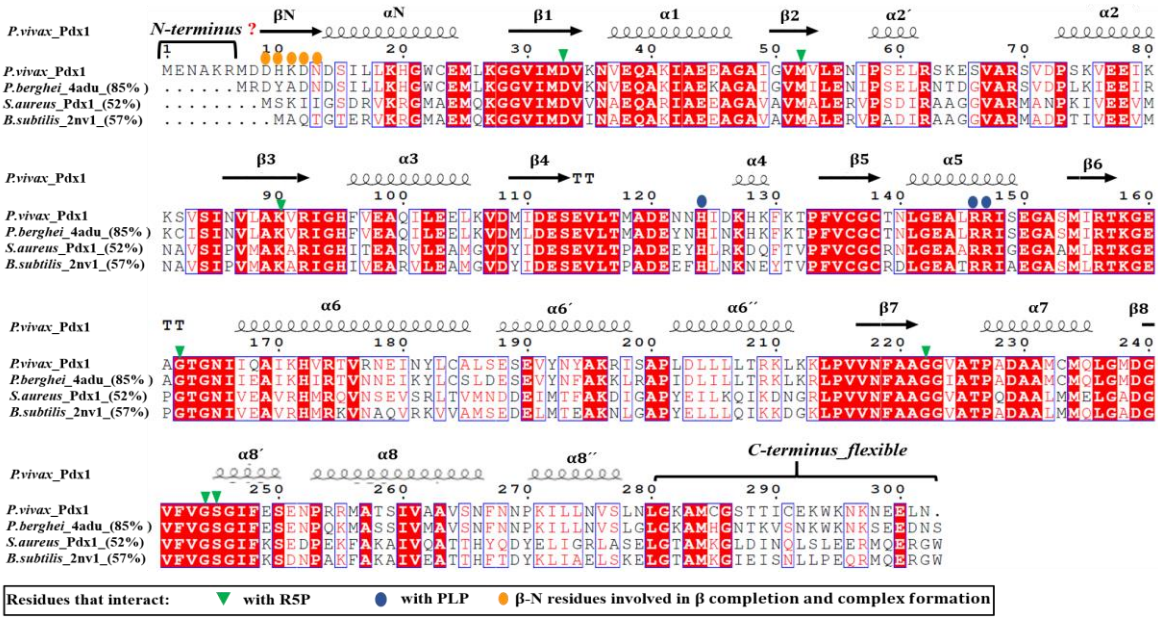


Figure-51. Sequence alignment of the *P.vivax* and *S.aureus* Pdx1 with homologous sequences. Pdx1 sequence from *P.vivax* (A5K247), *P.berghei* (Q4Z0E8), *S.aureus* (P60798), and *B.subtilis* (P37527). The red boxes with white letters show strict identity, red letters show similarity in the region between amino acid residues, and black letters show that regions are not conserved, with TT being a β turn. The green triangle labeled are the residues involved in binding of the substrate ribose 5-phosphate to the active site residue Lys-90 and the dark blue circles indicate the residues involved in product pyridoxal 5-phosphate binding and subsequent release. The orange-yellow dots mark the βN region, which interacts with the region (131–134) in Pdx2, involved in β-completion and complex stabilization in bacterial complexes. This β-completion is not reported in plasmodial species [69], but the elongated N-terminus in *Pv*Pdx1 (which is absent in other Plasmodial homologs) may have a role in complex stabilization and need further structural elucidation. The sequence alignment was prepared by Espript3 [96, 105].

Aiding to the sequence comparison, in recent years, the number of protein structures in the structure databases is expanding very rapidly. Therefore, the effectiveness and importance of comparison providing the structural and functional information about novel proteins are increasing. One of these methods of artificial intelligence is the measurement of the root-mean-square-deviation (RMSD) of the target structure with the reported data, where the root-mean-square distance between the corresponding residues in the structures is calculated. Using PyMol [172] and UCSF Chimera [102] superposition of the *Sa*Pdx1 structure to the three aligned homologous structures i.e. *B.subtilis* (Pdb code: 2NV1), *P.berghei* (Pdb code; 4ADU) and *Pv*Pdx1 were

performed, indicating corresponding RMSD values of 0.385, 0.647 and 0.611 Å for Cα atoms, respectively. Further, the superposition of the *Pv*Pdx1 structure was also performed to the homologous structures of *P.berghei* (Pdb code; 4ADU) and *B.subtilis* (Pdb code: 2NV1) revealed the corresponding Cα atoms RMSD values of 0.304, 0.684 Å respectively **(Figure 52)**.

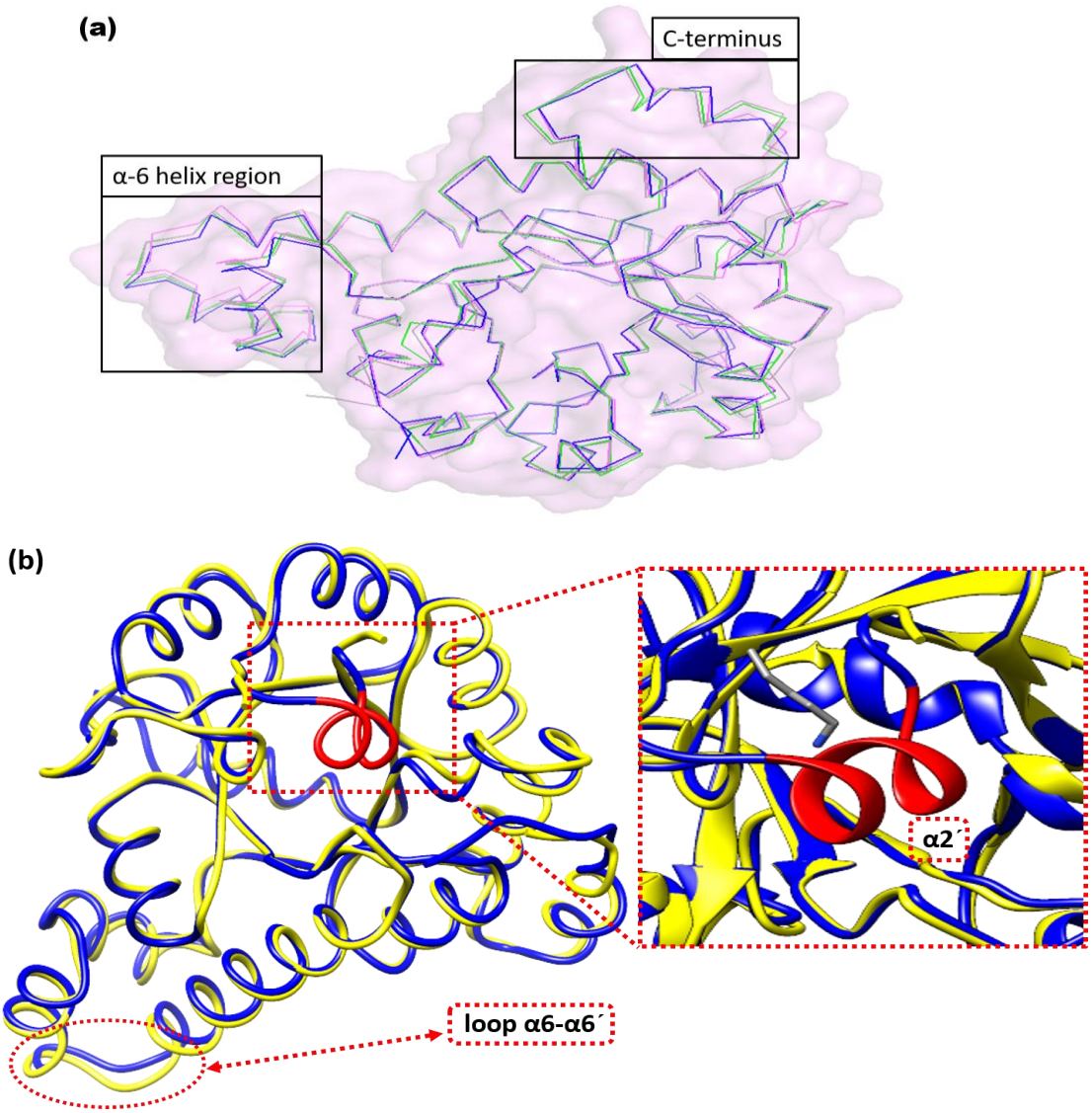


Figure-52. Structure superposition and RMS deviation of Pdx1 from *P.vivax* and *S.aureus*.
(a) Detailed structural superposition of Pdx1 from *P.vivax* (green) *P.berghei* (light grey), *S.aureus* (blue), and *B.subtilis* (purple) shown in ribbon representation, performed by PyMol [172]. Superposition results indicate the main chain RMSD in α6 with its insertion regions and also in the C-terminus region. **(b)** Structure superposition of *Sa*Pdx1 and *Pv*Pdx1, performed by UCSF Chimera [102], indicating loop α6-α6´ and C-terminus regions are showing notable deviation. The helix α2´ (red), which controls the entry of substrate (R5P) to the active site (Lys-90: stick) is ordered in the *S.aureus* Pdx1 structure (blue), whereas disordered in the *P.vivax* Pdx1 (yellow).

The superposition studies indicate that the superimposed structures are not much different spatially except for SaPdx1 in the helix α6 and C-terminus regions. Further, the detailed RMS deviations for the whole sequence of a monomer (chain-A) were calculated by software NOC (version 3.01). The full sequence comparison indicates that maximum deviation occurs at the C-terminus. The loop regions α6-α6′ (important in dodecamer interactions) and loop β8-α8′, as well as the N-terminus region, are also deviating. Further details of the whole chain and the amino acids showing the highest RMS deviations based on Cα are given in **figure 53**.

Table 22: Secondary structure β-strand content comparison of Pdx1 *P.vivax* and *S.aureus.*

No.	**Start**	**End**	**No. of residues**	**Edge**	**Sequence**
***P.vivax* Pdx1 β- strands**					
1	Val30	Val34	5	No	VIMDV
2	Gly50	Val53	4	No	GVMV
3	Asn86	Arg92	7	No	NVLAKVR
4	Met109	Ser113	5	No	MIDES
5	Phe134	Cys138	5	No	FVCGC
6	Met154	Thr157	4	No	MIRT
7	Val217	Phe219	3	Yes	VNF
8	Val241	Val243	3	Yes	VFV
***S.aureus* Pdx1 β- strands**					
1	Val22	Val26	5	No	VIMDV
2	Ala42	Ala42	4	No	AVMA
3	Val79	Ala83	5	No	VMAKA
4	Tyr101	Glu104	4	No	YIDE
5	Phe126	Cys130	5	No	FVCGC
6	Met146	Thr149	4	No	MLRT
7	Val209	Phe211	3	Yes	VNF
8	Val233	Val235	3	Yes	VFV

Table 22 shows that both structures have similar β-strands architecture throughout the Pdx1 β-barrel. The only difference is β3, which is 2 residues longer in *Pv*Pdx1. Noting this region contains the active site residue K90 in *Pv*Pdx1 (K82 in *Sa*Pdx1). Whereas table 23 indicates that both Pdx1 structures have very much conserved helices in terms of the number and sequence of amino acids. Interestingly, the half appeared helix αN in *Pv*Pdx1 contains a similar number of amino acids compared to the complete helix in *Sa*Pdx1. This shows that the lengthier *Pv*Pdx1 N-terminus need to be explored further in terms of structure and function.

Table 23: Secondary structure α-helices content comparison of Pdx1 *P.vivax* and *S.aureus.*

α- helices in *P.vivax* Pdx1 structure								
No	start	End	type	No of residue	Length	Unit rise	Res/turn	Sequence
1	Lys19	Leu26	H	8	13.05	1.46	3.72	KHGWCEML
2	Val37	Ala46	H	10	15.36	1.49	3.59	VEQAKIAEEA
3	Pro73	Lys80	H	8	12.76	1.51	3.69	PSKVEEIK
4	Phe96	Leu105	H	10	15.51	1.50	3.55	FVEAQILEEL
5	Lys127	Lys129	G	3	-	-	-	KHK
6	Leu141	Ser149	H	9	14.02	1.32	4.20	LGEALRRIS
7	Ile167	Ala185	H	19	29.03	1.51	3.64	IQAIKHVRTVRNEINYLCA
8	Glu188	Glu190	G	3	-	-	-	ESE
9	Val191	Arg197	H	7	10.74	1.44	3.62	VYNYAKR
10	Leu202	Leu211	H	10	15.10	1.47	3.59	LDLLLLTRKL
11	Pro227	Gln235	H	9	13.96	1.49	3.64	PADAAMCMQ
12	Gly246	Glu249	H	4	6.71	1.52	3.75	GIFE
13	Pro253	Ser265	H	13	19.73	1.49	3.64	PRRMATSIVAAVS
14	Pro270	Ser277	H	8	12.47	1.49	3.77	PKILLNVS
α- helices in *S.aureus* Pdx1 structure								
No	start	End	type	No of residue	Length	Unit rise	Res/turn	Sequence
1	Asp8	Glu16	H	9	14.03	1.49	3.71	DRVKRGMAE
2	Ala29	Ala38	H	10	15.41	1.49	3.64	AEQARIAEEA
α2′	Pro50	Ala56	H	7	10.90	1.49	3.64	PSDIRAA
3	Pro65	Met72	H	8	12.46	1.49	3.64	PKIVEEVM
4	Ile88	Ala96	H	9	14.02	1.50	3.64	ITEARVLEA
5	Lys119	Gln121	G	3	-	-	-	KDQ
6	Leu133	Gly141	H	9	13.59	1.44	3.65	LGEAARRIG
7	Val159	Thr176	H	18	27.47	1.49	3.66	VEAVRHMRQVNSEVSRLT
8	Asp180	Glu182	G	3	-	-	-	DDE
9	Ile183	Ile190	H	8	12.12	1.47	3.72	IMTFAKDI
10	Tyr194	Asn203	H	10	15.62	1.52	3.59	YEILKQIKDN
11	Pro219	Glu227	H	9	14.15	1.51	3.60	PQDAALMME
12	Ser237	Phe240	H	4	6.00	1.39	4.17	SGIF
13	Pro245	Thr257	H	13	19.81	1.49	3.62	PEKFAKAIVQATT
14	Tyr262	Leu268	H	7	10.85	1.48	3.61	YELIGRL

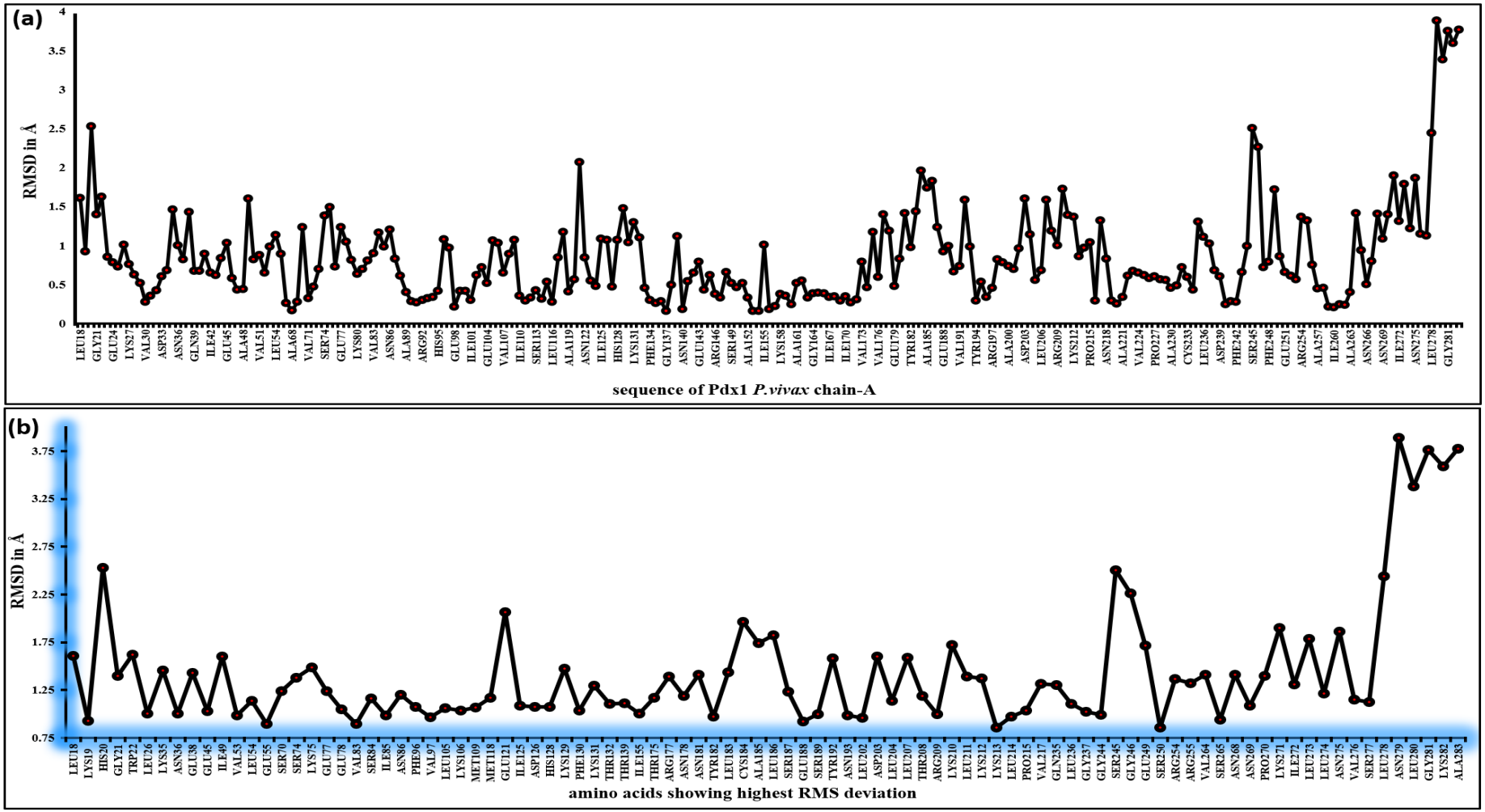


Figure-53. Whole chain RMSD of *P.vivax* versus *S.aureus* Pdx1
(a) Main chain RMSD (Å) of *P.vivax* Pdx1 chain-A vs *S.aureus* Pdx1 chain-A; **(b)** amino acids presenting the highest RMSD values after comparing both Pdx1 structures, indicating both terminal regions have the highest deviation.

5.0. Discussion

Despite continuous efforts by the WHO, malaria caused by a protozoan pathogen Plasmodium still exists as an endemic disease in Africa, South America, and Asia [3]. Drug resistance for *P.falciparum, P.vivax,* and *P.malariae* is recognized mainly as a result of long-term overuse of antimalarial antibiotics [3, 178]. Further, Methicillin-resistant *Staphylococcus aureus* (MRSA) is a major cause of hospitalized associated infections worldwide [19], constitutes an important public health challenge. This necessitates the identification of novel drug targets common at both evolutionary levels utilizing one shared approach.

Genome investigations indicate that most organisms, including bacteria, fungi, plants, and protozoans utilize class-I glutamine amidotransferase (GATase), which is an R5P-dependent pathway enzyme [103, 108], involved in the synthesis of pyridoxal 5-phosphate (PLP). The R5P-dependent pathway and its genes *pdx1* and *pdx2* are reported in several pathogenic organisms, such as eukaryotic protozoan parasites Plasmodium [108] and *Toxoplasma gondii* [179], unicellular eukaryotes yeasts i.e. *Saccharomyces cerevisiae*, *Candida utilis* [66] and bacteria, like *Mycobacterium tuberculosis* [180] and *B.subtilis* [90]. The disruption of the pdx gene in *M.tuberculosis* was reported lethal for the bacterium, indicating its importance for the survival of the organism [180]. Moreover, the functional *pdx1* and *pdx2* genes have been reported from *P.falciparum* expressing during their different host blood erythrocyte stages [103, 108].

In the class-I glutamine amidotransferase pathway, PLP (vitamin B_6) is synthesized by the PLP synthase complex consisting of the synthase (Pdx1) subunit and the glutaminase (Pdx2) subunit. The glutaminase subunit produces ammonia by utilizing its substrate glutamine. The nascent produced ammonia subsequently passes through the hydrophobic ammonia tunnel. This hydrophobic tunnel is present at the interface of both proteins, which ensure the safe passage of ammonia to the Pdx1 active site. Where this ammonia in the presence of intermediates from sugars R5P and G3P yield product PLP [103]. In this context, it is important to understand the structure and dynamic assembly of the Pdx complex, as well as its individual components (Pdx1 and Pdx2), which assemble to form the gigantic 12:12 Pdx complex.

Interestingly, Pdx1 and Pdx2 require each other for activation and substrates utilization, highlighting the ammonia tunnel entrance as an interesting target region for drug design and drug discovery investigations due to its presence at the protein-protein (Pdx1–Pdx2) interface. Further, both Pdx1 and Pdx2 are expressed throughout the intra-erythrocytic stages of Plasmodium, in the cytosol. They are able to interact and form the distinct Pdx complex, thereby generating vitamin B_6 during all developmental stages of the parasite. The absence of this metabolic pathway in humans and its essential presence in Plasmodium and bacteria [180], brands this pathway an interesting target for novel antimicrobial drugs [181].

5.1. Pdx1 oligomeric states and PO_4 binding sites role in dodecamer formation

It has been reported that the prokaryotic (*B.subtilis*) Pdx1 exists in an equilibrium of hexamers and dodecamers in solution [90]. For *G.stearothermophilus*, Pdx1 analytical ultra-centrifugation (AUC) studies showed a hexamer–dodecamer in solution with dodecameric form present in the crystalline form [97]. Further, the eukaryote *S.cerevisiae* Pdx1 (*Sc*Pdx1) is reported as a hexamer in solution as well as in crystalline state [99] whereas, the eukaryotic *P.berghei* Pdx1 is reported in a dodecameric form in solution as well as in crystalline state [69]. However, in contrast to the bacterial homologs and the eukaryote *S.cerevisiae Sc*Pdx1, the plasmodial *Pb* and *Pf* Pdx1 were reportedly forming higher-order oligomers/fibers investigated in SEC and EM studies [69]. It was assumed that the non-conserved and flexible C-terminus of the plasmodial Pdx1, which is absent in other homologs, is prone to degradation and presumably also involved in the fiber formation [69]. Therefore, plasmodial Pdx proteins were suggested as unattractive targets for structural studies, and further drug discovery investigations were aborted.

Difference in bacterial and plasmodial Pdx proteins sequences and also the distinct N and C-terminus regions in the sequence of *P.vivax* Pdx1 from other plasmodial homologs, I opted *P.vivax* Pdx proteins for structure comparison with *S.aureus* Pdx proteins, leading to future drug discovery studies. In the first part of my thesis, *P.vivax* Pdx proteins were investigated applying complementary biophysical and bioanalytical techniques and compared the results to its homologous plasmodial Pdx proteins. *P.vivax* Pdx1 was investigated by SEC, DLS, and mass photometry studies, presenting a long-term solution stable dodecameric state. Further, SAXS data and

EM investigations are revealing a low-resolution dodecameric structure of the *Pv*Pdx1 in solution. The sequence alignment given in **figure 51** indicates that helices α6, α6′, and α6′′, which stabilize the dodecamer, are highly conserved among the plasmodial species and significant variations are only observed at the N- and C-termini of the sequence [166]. Subsequently, *S.aureus* Pdx1 was also investigated by the above-mentioned biophysical and bioanalytical techniques and revealing a hexameric form in solution. The sequence comparison studies of the Pdx1 subunit reveal that structurally important regions are much conserved in both species. This conditioned further detailed comparative structural investigations, for the reason of their different biologically relevant oligomeric solution states.

Thus, Pdx1 from both species has proceeded further for comparative structural investigations by X-ray crystallography. Crystal structure results indicate *P.vivax* Pdx1 in the dodecameric state whereas *S.aureus* Pdx1 oligomeric state, in the preliminary 3D structural data, was a bit puzzling. Therefore, both the structures were compared and investigated comprehensively. At first glance, the Pdx1 protein starting helix (helix αN) has appeared differently. Half of the helix αN appeared for *Pv*Pdx1 and full-length αN is solved for *Sa*Pdx1, both with a similar spatial orientation to the reported data. The sequence of helix αN is slightly lengthier in *P.vivax* than observed in other plasmodial cognates and bacterial homologs, with a role in the interplay between Pdx1 and Pdx2 as well as stability of the dodecameric Pdx1 and the Pdx complex [182].

To summarize both Pdx1 structures comparatively, there are two phosphate-binding sites on each Pdx1 monomer, involved in the substrate (R5P) accommodation and product (PLP) formation, referred to as P1 and P2 sites, respectively. In the absence of substrate and product, buffer salt anions chloride ions [90] or sulfate ions [97] were observed and reported at these positions. Position P1 is located on the C-terminal side of the $(\beta\alpha)_8$ barrel. It comprises a characteristic loop for phosphate binding [97] which constitutes the amino acid residues ^{154}Gly-^{156}Gly in *S.aureus* (GTG 162-164 in *P.vivax*). This loop is close to the R5P binding residue (active site) Lys-82 in *S.aurues* Pdx1 (Lys-90 in *P.vivax*) lined with Asp-25 (Asp-33 in *P.vivax*). It resides chloride ion [69], which is linked to glycerol moiety in *P.vivax* Pdx1 through non-covalent interaction. Whereas in *S.aureus* Pdx1, the chloride ion is

linked to the PO_4 ion through ionic interaction. The second site (P2) is located at the interface of the hexameric rings and is involved in dodecamer stabilization [90]. It is characterized by residues His-116 (His-124 in *P.vivax* Pdx1), Arg-138, and Arg-139 (Arg-146, Arg-147 *P.vivax* Pdx1). Which are linked through PO_4 ion to Lys-188 (Lys-196 in *P.vivax*) of the monomer from opposing hexamer. Both P1 and P2 sites are spatially located at a distance of approx. 20 Å.

5.2. *P.vivax* Pdx1 and Pdx complex rational drug targets

Since, higher-order oligomers and fiber formation are reported from the *P.falciparum* Pdx complex upon storage at 4°C, as observed by analytical SEC and EM analysis [69]. Whereas the *P.vivax* Pdx complex investigated in the thesis by applying SEC and DLS supported by SAXS and EM studies indicate that the complex is stable in solution without observing any higher-order oligomers over time. Further, no significant concentration-dependent oligomerization or aggregation was identified for Pdx1 and the Pdx complex at concentrations ranging from 1 to 30 mg mL^{-1}. The solution stable *P.vivax* proteins could be utilized as potential drug targets for antimalarial drug development. However, the *S.aureus* Pdx complex was showing different behavior in solution, studied by SEC and DLS and established by SAXS investigation, indicate an unusual tetrameric Pdx complex. Further, the EM investigation of *S.aureus* Pdx1 and Pdx complex do not show pronounced results at different concentrations, indicating stability issues with staining and proteins disintegration into smaller subunits.

5.3. Bacterial *Sa*Pdx1 novel helix α2' following eukaryotic architecture

Helix α2' is a very crucial and important structural element of Pdx1, located very close to its catalytic center. When substrate R5P enters and binds covalently to the active site residue lysine (*S.aureus* Lys-82), helix α2' closes the active site to protect the substrate intermediates from the solvent and prepare the enzyme for further biological steps and also in neutral NH_3 transport from the Pdx2 subunit. The helix is found disordered in many reported native Pdx1 structures [69, 90, 92].

Earlier structural investigations of bacterial Pdx proteins showed that the disordered helix α2' of Pdx1 was oriented and ordered after interaction with Pdx2 in Pdx complex formation from *T.maritima* (Pdb code: 2ISS) [92] and *B.subtilis* (Pdb code: 2NV2) [90]. Thus, it was assumed that the binding of Pdx2 induces an allosteric

effect on Pdx1 for enhanced R5P binding affinity [87]. Similar allosteric induction was also observed in plasmodial PLP synthase, where the helix α2' is disordered in the native Pdx1. Though, it appeared in the Pdx1-Pdx2 complex but at a position different to the reported bacterial Pdx complexes and do not cover the Pdx1 phosphate binding site (P1). Further, it was also found that the addition of R5P also induces this helix ordering in plasmodial Pdx1 [69], with its position closer to what was observed in the bacterial Pdx complexes. In another study, on yeast *Saccharomyces cerevisiae* hexameric Pdx1 native crystal structures (Pdb codes 3O06) [91] and 3FEM [99], showed an ordered α2' helix in a position similar to *Pb*Pdx1-R5P **(Figure 54)** [69]. By investigating the appearance and three-dimensional spatial position of helix α2' in the *S.aureus* Pdx1 structure reveals that it follows the *S.cerevisiae* Pdx1 and *Pb*Pdx1-R5P orientation, as shown in **figure 54**. This orientation can be called an "open position", which allows the entry of NH_3 to the active site through the hydrophobic ammonia tunnel.

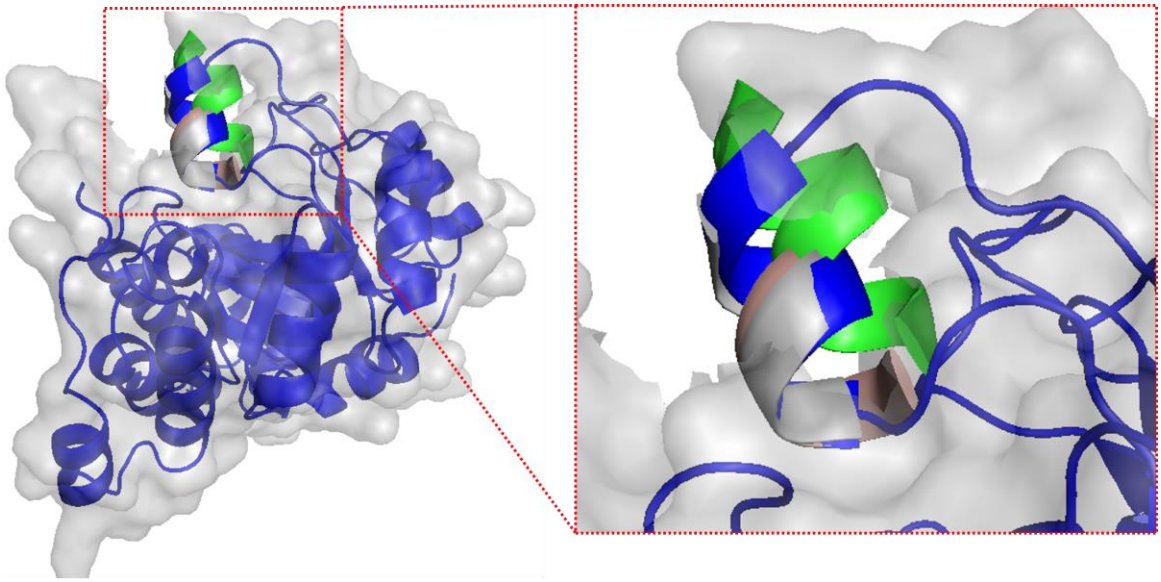


Figure-54. The relative position and orientation of helix α2' in the Pdx1 structures.

Green: *B.subtilis* Pdx1 subunit from Pdx complex (the helix do not appear in native *Bs*Pdx1), Light brown: Pdx1 from *S.cerevisiae*, Light grey: Pdx1 from *P.berghei*-R5P adduct and Blue is Pdx1 from *S.aurues*, reported in this study. The relative positions indicate that in native Pdx1 and Pdx1 after complex formation with Pdx2 (*B.subtilis*) the helix α2' is positioned in different orientations.

5.4. Pdx1 C-terminus: A key player in PLP catalysis and oligomerization

The functionally conserved but less sequence conserved C-terminus of Pdx1 plays a major role in the PLP synthase catalysis and also in the oligomerization of the Pdx1 hexameric architecture [87, 99]. Because of the flexible nature [90, 92, 99], the 3D structure of the full C-terminus is not yet resolved. To date, the reported data indicate that the C-terminus forms an extended loop in position adjacent to the loop α2′-α2 of the interface monomer, establishing inter-chain contacts for Pdx1 hexamer formation. In one study, the 23 C-terminal loop residues of Pdx1 from *P.falciparum* and *P.berghei* were deleted. This results in loss of the PLP synthase activity, but with intact hexamer and dodecamer, indicating its role in enzyme catalysis. In the next experiment, the sequence was mutated upstream with 3 more residues deletion (end residues of C-terminal helix α8") and the results show that Pdx1 is in monomer form with dimer and trimer molecules in equilibrium. The deletion data corroborate the C-terminal helix α8" critical role in oligomerization [99, 166].

However, the obtained *P.vivax* and *S.aureus* Pdx1 3D structures indicate that the C-terminal helix α8" is not involved in direct bonding interactions between the protomers. Instead, the mid part of the C-terminus establishes contacts to the interface chain, suggesting helix α8" may have an indirect role in the oligomerization of Pdx1. In both the structures, the flexible C-terminus loop runs in parallel and is sandwiched between the loop α2′-α2 and loop β6-α6, establishing a network of hydrogen bonds and connect interface protomers for oligomerization. The loop β6-α6 contacts indicate stabilization of the flexible C-terminus and prepare it for proper bonding to the protomer loop α2′-α2. Further, the structurally important residue Arg-61 in *S.aureus* (Arg-69 in *P.vivax*), a highly conserved residue in the α2′-α2 loop, is involved in establishing contact between the C-terminus and loop (β6-α6) at the adjacent protomer subunit.

5.5. Pdx2 occupies the Pdx1 dodecamer randomly

For both plasmodial and bacterial PLP synthases, Pdx2 is active only in the presence of dodecameric Pdx1 [68, 103]. The interaction between the two proteins (Pdx1 and Pdx2) is mediated by the conserved helix αN of Pdx1 [94]. After analyzing the bacterial *B.subtilis* Pdx complex, it has been observed that the complex requires the catalytically inactive Pdx2 in the catalytic triad (residue Histidine) for Pdx1 and

Pdx2 interactions strengthening, and the complex stabilization [90]. Further, the Plasmodium (*P.berghei*) Pdx proteins, with active and inactive Pdx2 variants, were investigated by electron microscopy. The results showed that a few Pdx2 subunits from the wild-type variants occupy the Pdx1 dodecamer. Whereas, the inactive $Pdx2^{H196N}$ subunits fully occupied *Pb*Pdx1 dodecameric ring and resulted in saturated Pdx1-Pdx2 complex particles [69].

In *P.falciparum*, Pdx2 was reported previously to be monomeric in solution and a dimer in the crystal lattice [103]. In contrast, the *P.vivax* Pdx2 is predominantly monomeric in solution. However, a fraction of oligomers was identified in the solution in DLS and SAXS experiments. The CD results indicate approximately 64% β-sheets and coils in *Pv*Pdx2, thereby showing that Pdx2 is prone to some oligomerization as increased β-sheet content increases the potential for oligomerization [183]. TANGO, a statistical mechanics algorithm based on the physicochemical principles of β-sheet formation to detect protein oligomerization [184], suggest that the regions 4–10, 30–35, 84–94, 182–186, and 193-196 of *Pv*Pdx2, which form β-sheets, support oligomerization. Whereas in *Sa*Pdx2 fewer regions are involved in β-sheet formation tending to oligomerize. These are 3-12, 72-82, and 152-155, indicating different sequence positions from *Pv*Pdx2. Further, the *S.aureus* Pdx2 was found only in the monomeric state, investigated by SEC, DLS, and SAXS.

5.6. Failures in the Pdx complex structure determination

Following the reported PLP synthase data, many manual crystallization attempts, as well as high-throughput screening, were performed for *P.vivax* and *S.aureus* Pdx complexes. The resultant fragile crystals from *P.vivax* Pdx complex diffracted only to low resolution and failed crystallization for *S.aureus* Pdx complex. The purified Pdx complex was further investigated by electron microscopy, indicating an uneven distribution of the Pdx2 subunit at the Pdx1 dodecameric core in the *P.vivax* Pdx complex. The possible reason is after utilizing its substrate (Glutamine), Pdx2 detaches from Pdx1 and resulted in a poly-dispersed solution of Pdx complex molecules with a varying number of bound Pdx2 molecules. This also results in alterations of the symmetry of Pdx complex molecules. Such a solution is not a good

starting point for crystallization; therefore, further crystallization attempts were aborted.

Hence, It has been reported that the Pdx2 active site mutation at H196N resulted in a fully decorated saturated Pdx1-Pdx2 complex, which resulted in a monodisperse solution [69]. But the complex subsequently loses its inhibitors binding ability [68]. In another study on *S.cerevisiae* type-1 GATase (Imidazole Glycerol Phosphate Synthase), glutamine analog acivicin was detected in a covalently bound state to the catalytic triad residue (Cys-83) [185]. Therefore, active site mutation was avoided in this thesis and instead a few potent Pdx2 bound inhibitors were utilized for Pdx complex stabilization and inhibition. The kinetic assays indicate inhibition of the Pdx complex but also failed to crystallize.

6. Conclusion and Future Outlook

Due to the absence of the ribose 5-phosphate (R5P) dependent pathway enzymes in humans and their presence in human parasites and pathogens, the PLP synthase (Pdx1-Pdx2 complex) enzyme has emerged recently as an attractive drug target. PLP synthase, a GATase enzyme, function as a glutamine amidotransferase and produce pyridoxal 5-phosphate (PLP), an important vitamin needed for the survival of different human parasites and pathogens. Pdx complex is composed of two subunits called Pdx1 and Pdx2, where Pdx2 functions as the glutaminase domain to metabolize substrate glutamine into neutral ammonia. Whereas, Pdx1 is the synthase domain which utilizes the substrates ribose 5-phosphate (R5P), glyceraldehyde 3-phosphate (G3P), and ammonia. The reaction mechanism between the two enzymes is cooperative, with the ammonia produced by Pdx2 is transported to the active site of the synthase (Pdx1) domain through a hydrophobic ammonia tunnel. The NH_3 is condensed with R5P at Pdx1 active site, forming a unique intermediate (I_{320}) which is afterward converted into PLP by reacting with G3P. Since Pdx1 and Pdx2 depend on each other for activation and also for their subsequent enzymatic reactions linking through ammonia tunnel. This hydrophobic tunnel is an important vulnerable site for inhibitors design, based on the principle of blocking the tunnel.

As the disease-causing human pathogens belong to different evolutionary levels and the reported Pdx complex has significant differences reported in different prokaryote and eukaryote species. Based on these differences, the PLP synthase enzymes from *Plasmodium vivax* and *Staphylococcus aureus* were investigated in detail applying different biophysical approaches like DLS, SAXS, mass photometry (MP), and nanoparticles tracking analysis (NTA). Investigations were also supported by electron microscopy analysis of the Pdx proteins. Subsequently, the selected Pdx proteins 3D structures were also analyzed. Since the Pdx1 enzyme subunit is the major player in the PLP synthesis, performing a complex series of reactions. There are around 20 or more enzymatic steps with each having a structurally unique intermediate involved in the PLP synthesis, with most of them occur at the Pdx1 barrel. Therefore, more focused and detailed comparative structural investigations of Pdx1 are required.

In the first part of my thesis, Pdx1, Pdx2, and Pdx1-Pdx2 complex from *P.vivax* were investigated applying biophysical techniques like DLS and SAXS, indicating *Pv*Pdx1 in solution stable cylindrical dodecameric form. Mass photometry and EM investigations also indicate its monodisperse dodecameric form aiding its stability confirmation. *Pv*Pdx complex investigations via SEC, DLS, and SAXS, supported by EM studies, also indicate its solution stable form. Further, the *Pv*Pdx1 crystal structure shows many common structural features to the reported Pdx1 data. The Pdx1 N-terminal helix (αN) appeared half in *Pv*Pdx1, reportedly present at the interface of proteins (Pdx1 and Pdx2) and is involved in the Pdx1-Pdx2 assembly. Further, the unique elongated N-terminus sequence appeared only in *P.vivax* Pdx1, present before the helix αN, which needs to be investigated in detail regarding its potential role in Pdx complex stabilization and catalysis.

The counterpart Pdx2 oligomerization is reported for the first time and that its oligomerization is dependent on time and its concentration, established by DLS, NTA, and EM studies. Further, the *Pv*Pdx2 concentration-dependent oligomerization is reversible upon interacting with the counterpart Pdx1 dodecamer, studied by time-resolved DLS, and supported by NTA experimentation. The reversible oligomerization of *Pv*Pdx2 toward complex formation with Pdx1 was also confirmed by electron microscopy investigations. The EM analysis displayed the Pdx complex formation after oligomeric Pdx2 interacted with the dodecameric Pdx1.

The second part of my thesis includes detailed investigations of the *S.aureus* Pdx1, Pdx2, and Pdx complex, providing comparative information about the plasmodial Pdx proteins. The investigations aimed to find similarities and differences in the prokaryotic and eukaryotic Pdx proteins to design a common drug in the future. The *S.aureus* monomeric Pdx2 is stable and following the other reported homologous prokaryotic Pdx2 proteins behaviour in solution in terms of stability and solution states. Surprisingly, *Sa*Pdx1 is found in a hexameric solution state without any equilibrium dodecameric form. *Sa*Pdx1 is the first reported prokaryotic Pdx1 found as a hexamer in solution without a dodecamer form at equilibrium.

Further, in the *Sa*Pdx1 structure the starting residues of helix α6′ (important in dodecamer formation) showed higher deviation from its homologous bacterial and *P.vivax* Pdx1 structures which results in a shorter helix α6′, with less number of

inter-hexameric contacts. Next, the helix α6 and α6′ residues Asn-178, Asn-181 in *Pv*Pdx1 that are involved in dodecamer formation are mutated to Ser-170 and Ser-173 in *Sa*Pdx1 with no inter-chain bonding. Similarly, Asn-169 (Arg-177 in *P.vivax*) and Met-184 (Tyr192 in *P.vivax*) appeared in *Sa*Pdx1 with loss in inter-chain (inter-hexamer) salt bridging. Interestingly, the presence of phosphate ion at the hexamer interface in *Sa*Pdx1, in the proximity of the helices α6′ and α5, connect interfacing monomers through hydrogen bonding. Noting that these PO_4 ions result from the crystallization condition of *Sa*Pdx1, including different metals phosphate and sulfate salts.

S.aureus Pdx complex investigated here shows different solution behavior in biophysical studies like SEC and DLS, with subsequent X-ray solution scattering (SAXS) results. Generally, dodecameric Pdx1 after interaction with Pdx2 forms a 24meric Pdx1-Pdx2 complex. However, the *Sa*Pdx complex is displaying a tetrameric form in SAXS data with a 1:1 ratio of Pdx1 and Pdx2. Further, the Pdx complex from hexameric Pdx1 and monomeric Pdx2 is not reported in any in-vitro or in-vivo studies, until now. Thus, *Sa*Pdx complex assembly **(Figure 55)** investigation is a wise approach, with potential new information providing about the PLP synthase complex machinery structure and function.

In summary, compared to other plasmodial homologs, Pdx1 and the Pdx complex from *P.vivax* are more stable in solution than their homologs. Therefore, the stable *P.vivax* Pdx complex structure characterization in terms of the PLP biosynthetic mechanism utilizing X-ray crystallography with substrate analog complexes provides snapshots of different catalytic steps would be a good future strategy in this project. Further, after analyzing Pdx1 from *P.vivax* and *S.aurues* comparatively by applying different biophysical and bioanalytical techniques, indicates their respective stable solution states. Their comparative crystal structures demonstrate that the functionally important common structural features need to be considered and investigate as drug target regions for future drug design studies.

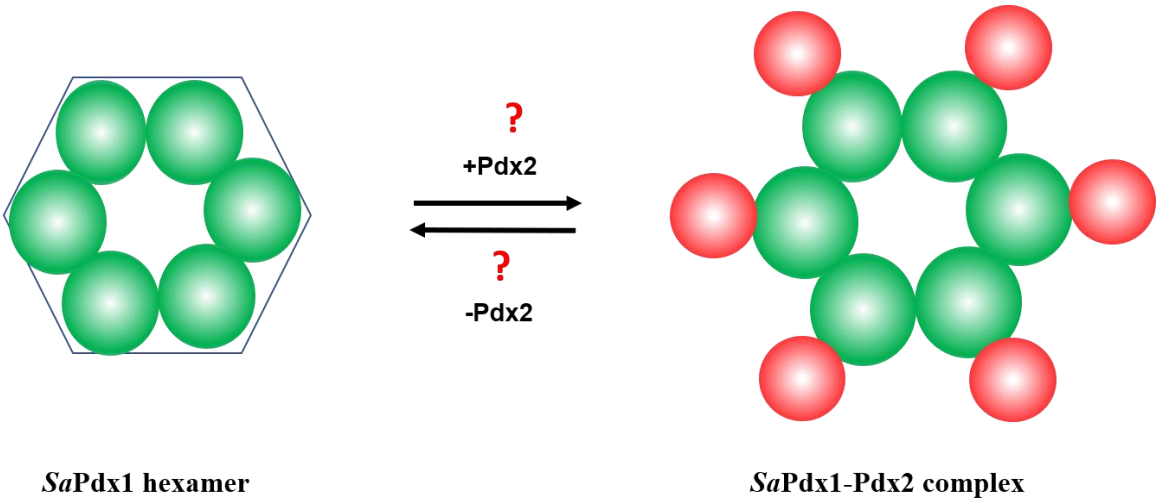


Figure-55. Schematic representation of the *S.aureus* Pdx1 and Pdx complex.

*Sa*Pdx1 schematic representation, where each circle (green) is a Pdx1 monomer, arranged in hexamer form in solution and after reacting with Pdx2 (red) forming a potential Pdx1-Pdx2 complex.

References

1. Clancy, C.J. and M.H. Nguyen. *Buying time: The AMR Action Fund and the state of antibiotic development in the United States 2020*. in Open Forum Infectious Diseases. Oxford University Press US, 2020.
2. Organization, W.H., *World malaria report 2019. Geneva:* World Health Organization; *2019*. 2020.
3. Amadi, A., I. Ezeonu, and O. Akoma, *Antimalarial Drug Resistance: An Existential Burden for the Developing World.* Microbiology Research Journal International, 2019: p. 1-16.
4. Capanna, E., *Grassi versus Ross: who solved the riddle of malaria?* International Microbiology, 2006. **9**(1): p. 69-74.
5. Organization, W.H., *World malaria report 2020: 20 years of global progress and challenges.* 2020.
6. Bruce-Chwatt, L.J., C. Garrett-Jones, and B. Weitz, *Ten years' study (1955-64) of host selection by anopheline mosquitos.* Bulletin of the World Health Organization, 1966. **35**(3): p. 405.
7. Soulard, V., et al., *Plasmodium falciparum full life cycle and Plasmodium ovale liver stages in humanized mice.* Nature communications, 2015. **6**(1): p. 1-9.
8. Twohig, K.A., et al., *Growing evidence of Plasmodium vivax across malaria-endemic Africa.* PLoS neglected tropical diseases, 2019. **13**(1): p. e0007140.
9. Organization, W.H., *Control and elimination of Plasmodium vivax malaria: a technical brief*. World Health Organization,.2015.
10. WHO, I., *Guidelines for the treatment of malaria.* Geneva: World Health Organization Press, 2010.
11. Patel, T.S., et al., *Design and synthesis of leucine-linked quinazoline-4 (3H)-one-sulphonamide molecules distorting malarial reductase activity in the folate pathway.* Archiv der Pharmazie, 2019. **352**(9): p. 1900099.
12. Clifton, C., *Large-scale production of penicillin.* Science, 1943. **98**(2533): p. 69-70.
13. Fleming, A., *Penicillin: Nobel lecture, December 11*. 1945.
14. San Millán, Á., *Dodging magic bullets: The evolution of bacterial antibiotic resistance.* Mètode Science Studies Journal, 2020. **10**: p. 207-211.

15. O'Neill, J., *Securing new drugs for future generations: the pipeline of antibiotics.* Wellcome Library, 2015.
16. Van Boeckel, T.P., et al., *Global antibiotic consumption 2000 to 2010: an analysis of national pharmaceutical sales data.* The Lancet infectious diseases, 2014. **14**(8): p. 742-750.
17. Heinemann, J.A., *How antibiotics cause antibiotic resistance.* Drug discovery today, 1999. **4**(2): p. 72-79.
18. D'Costa, V.M., et al., *Sampling the antibiotic resistome.* Science, 2006. **311**(5759): p. 374-377.
19. ECDC, E., *The bacterial challenge: time to react European Centre for Disease Prevention and Control European Medicines Agency.* 2009.
20. Köck, R., et al., *Methicillin-resistant Staphylococcus aureus (MRSA): burden of disease and control challenges in Europe.* Eurosurveillance, 2010. **15**(41): p. 19688.
21. Víquez-Molina, G., et al., *Virulence factor genes in Staphylococcus aureus isolated from diabetic foot soft tissue and bone infections.* The international journal of lower extremity wounds, 2018. **17**(1): p. 36-41.
22. Spellberg, B., J.G. Bartlett, and D.N. Gilbert, *The future of antibiotics and resistance.* New England Journal of Medicine, 2013. **368**(4): p. 299-302.
23. Lewis, K., *Recover the lost art of drug discovery.* Nature, 2012. **485**(7399): p. 439-440.
24. Begum, A., et al., *Staphylococcus aureus thiaminase II: oligomerization warrants proteolytic protection against serine proteases.* Acta Crystallographica Section D: Biological Crystallography, 2013. **69**(12): p. 2320-2329.
25. Ohdake, S., *Isolation of "Oryzanin"(Antineuritic Vitamin) from Rice-polishings.(First Report.).* Bulletin of the Agricultural Chemical Society of Japan, 1932. **8**(1-3): p. 11-46.
26. Gyürgy, P. and R.E. Eckardt, *Vitamin B 6 and skin lesions in rats.* Nature, 1939. **144**(3646): p. 512-512.
27. Harris, S.A. and K. Folkers, *Synthetic vitamin B6.* Science (Washington), 1939. **89**.

28. Perry, T., et al., *Evidence of GLP-1-mediated neuroprotection in an animal model of pyridoxine-induced peripheral sensory neuropathy.* Experimental neurology, 2007. **203**(2): p. 293-301.
29. Albin, R., et al., *Acute sensory neuropathy-neuronopathy from pyridoxine overdose.* Neurology, 1987. **37**(11): p. 1729-1732.
30. Percudani, R. and A. Peracchi, *A genomic overview of pyridoxal-phosphate-dependent enzymes.* EMBO reports, 2003. **4**(9): p. 850-854.
31. Mooney, S. and H. Hellmann, *Vitamin B6: Killing two birds with one stone?* Phytochemistry, 2010. **71**(5-6): p. 495-501.
32. Christen, P. and P.K. Mehta, *From cofactor to enzymes. The molecular evolution of pyridoxal-5'-phosphate-dependent enzymes.* The Chemical Record, 2001. **1**(6): p. 436-447.
33. Eliot, A.C. and J.F. Kirsch, *Pyridoxal phosphate enzymes: mechanistic, structural, and evolutionary considerations.* Annual review of biochemistry, 2004. **73**(1): p. 383-415.
34. Hanada, K., *Serine palmitoyltransferase, a key enzyme of sphingolipid metabolism.* Biochimica et Biophysica Acta (BBA)-Molecular and Cell Biology of Lipids, 2003. **1632**(1-3): p. 16-30.
35. Livanova, N., et al., *Pyridoxal 5"-Phosphate as a Catalytic and Conformational Cofactor of Muscle Glycogen Phosphorylase b.* Biochemistry (Moscow), 2002. **67**(10): p. 1089-1098.
36. Frey, P.A., *Radical mechanisms of enzymatic catalysis.* Annual review of biochemistry, 2001. **70**(1): p. 121-148.
37. He, X., G. Agnihotri, and H.-w. Liu, *Novel enzymatic mechanisms in carbohydrate metabolism.* Chemical reviews, 2000. **100**(12): p. 4615-4662.
38. Fitzpatrick, T.B., et al., *Two independent routes of de novo vitamin B6 biosynthesis: not that different after all.* Biochemical Journal, 2007. **407**(1): p. 1-13.
39. Rodrigues, M.J., et al., *Lysine relay mechanism coordinates intermediate transfer in vitamin B6 biosynthesis.* Nature chemical biology, 2017. **13**(3): p. 290-294.

40. Garrido-Franco, M., *Pyridoxine 5'-phosphate synthase: de novo synthesis of vitamin B6 and beyond.* Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 2003. **1647**(1-2): p. 92-97.
41. Banks, J. and D.E. Cane, *Biosynthesis of vitamin B6: direct identification of the product of the PdxA-catalyzed oxidation of 4-hydroxy-l-threonine-4-phosphate using electrospray ionization mass spectrometry.* Bioorganic & medicinal chemistry letters, 2004. **14**(7): p. 1633-1636.
42. Garrido Franco, M., et al., *Crystallization and preliminary X-ray crystallographic analysis of PdxJ, the pyridoxine 5'-phosphate synthesizing enzyme.* Acta Crystallographica Section D: Biological Crystallography, 2000. **56**(8): p. 1045-1048.
43. Sivaraman, J., et al., *Crystal structure of Escherichia coli PdxA, an enzyme involved in the pyridoxal phosphate biosynthesis pathway.* Journal of Biological Chemistry, 2003. **278**(44): p. 43682-43690.
44. Mukherjee, T., et al., *Pyridoxal phosphate: biosynthesis and catabolism.* Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 2011. **1814**(11): p. 1585-1596.
45. Zhao, G., et al., *Biochemical characterization of gapB-encoded erythrose 4-phosphate dehydrogenase of Escherichia coli K-12 and its possible role in pyridoxal 5'-phosphate biosynthesis.* Journal of Bacteriology, 1995. **177**(10): p. 2804-2812.
46. Drewke, C., et al., *4-O-phosphoryl-L-threonine, a substrate of the pdxC (serC) gene product involved in vitamin B6 biosynthesis.* FEBS letters, 1996. **390**(2): p. 179-182.
47. Zhao, G. and M.E. Winkler, *A novel alpha-ketoglutarate reductase activity of the serA-encoded 3-phosphoglycerate dehydrogenase of Escherichia coli K-12 and its possible implications for human 2-hydroxyglutaric aciduria.* Journal of bacteriology, 1996. **178**(1): p. 232-239.
48. Cane, D.E., et al., *Biosynthesis of vitamin B6: the oxidation of 4-(phosphohydroxy)-L-threonine by PdxA.* Journal of the American Chemical Society, 1998. **120**(8): p. 1936-1937.
49. Kuzuyama, T., et al., *Cloning and Characterization of 1-Deoxy-d-Xylulose 5-Phosphate Synthase fromStreptomyces sp. Strain CL190, Which Uses both the*

Mevalonate and Nonmevalonate Pathways for Isopentenyl Diphosphate Biosynthesis. Journal of Bacteriology, 2000. **182**(4): p. 891-897.

50. Cane, D.E., et al., *Biosynthesis of vitamin B6: enzymatic conversion of 1-deoxy-D-xylulose-5-phosphate to pyridoxol phosphate.* Journal of the American Chemical Society, 1999. **121**(33): p. 7722-7723.
51. Laber, B., et al., *Vitamin B6 biosynthesis: formation of pyridoxine 5'-phosphate from 4-(phosphohydroxy)-L-threonine and 1-deoxy-D-xylulose-5-phosphate by PdxA and PdxJ protein.* FEBS letters, 1999. **449**(1): p. 45-48.
52. Yeh, J.I., et al., *Multistate binding in pyridoxine 5 '-phosphate synthase: 1.96 Å crystal structure in complex with 1-deoxy-d-xylulose phosphate.* Biochemistry, 2002. **41**(39): p. 11649-11657.
53. Di Salvo, M., et al., *Expression, Purification, and Characterization of Recombinant Escherichia coli Pyridoxine 5'-Phosphate Oxidase.* Protein expression and purification, 1998. **13**(3): p. 349-356.
54. Zhao, G. and M.E. Winkler, *Kinetic limitation and cellular amount of pyridoxine (pyridoxamine) 5'-phosphate oxidase of Escherichia coli K-12.* Journal of Bacteriology, 1995. **177**(4): p. 883-891.
55. di Salvo, M.L., et al., *Structure and mechanism of Escherichia coli pyridoxine 5'-phosphate oxidase.* Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 2003. **1647**(1-2): p. 76-82.
56. Ehrenshaft, M., et al., *A highly conserved sequence is a novel gene involved in de novo vitamin B6 biosynthesis.* Proceedings of the National Academy of Sciences, 1999. **96**(16): p. 9374-9378.
57. Ehrenshaft, M. and M.E. Daub, *Isolation of PDX2, a second novel gene in the pyridoxine biosynthesis pathway of eukaryotes, archaebacteria, and a subset of eubacteria.* Journal of Bacteriology, 2001. **183**(11): p. 3383-3390.
58. Tazuya, K., et al., *Origin of the nitrogen atom of pyridoxine in Saccharomyces cerevisiae.* Biochimica et Biophysica Acta (BBA)-General Subjects, 1995. **1244**(1): p. 113-116.
59. Tanaka, K., et al., *Biosynthesis of pyridoxine: origin of the nitrogen atom of pyridoxine in microorganisms.* Journal of Nutritional Science and Vitaminology, 2000. **46**(2): p. 55-57.

60. Mittenhuber, G., *Phylogenetic analyses and comparative genomics of vitamin B6 (pyridoxine) and pyridoxal phosphate biosynthesis pathways.* Journal of Molecular Microbiology and Biotechnology, 2001. **3**(1): p. 1-20.
61. Belitsky, B.R., *Physical and enzymological interaction of Bacillus subtilis proteins required for de novo pyridoxal 5'-phosphate biosynthesis.* Journal of Bacteriology, 2004. **186**(4): p. 1191-1196.
62. Sakai, A., et al., *yaaD and yaaE are involved in vitamin B6 biosynthesis in Bacillus subtilis.* Journal of Bioscience and Bioengineering, 2002. **93**(3): p. 309-312.
63. Wetzel, D.K., et al., *Functional complementation between the PDX1 vitamin B6 biosynthetic gene of Cercospora nicotianae and pdxJ of Escherichia coli.* FEBS Letters, 2004. **564**(1-2): p. 143-146.
64. Ehrenshaft, M., et al., *Functional characterization of SOR1, a gene required for resistance to photosensitizing toxins in the fungus Cercospora nicotianae.* Current Genetics, 1999. **34**(6): p. 478-485.
65. Ehrenshaft, M., et al., *SOR1, a gene required for photosensitizer and singlet oxygen resistance in Cercospora fungi, is highly conserved in divergent organisms.* Molecular Cell, 1998. **1**(4): p. 603-609.
66. Zeidler, J., et al., *Biosynthesis of vitamin B6 in yeast. Incorporation pattern of trioses.* The Journal of Organic Chemistry, 2003. **68**(9): p. 3486-3493.
67. Burns, K.E., et al., *Reconstitution and biochemical characterization of a new pyridoxal-5 '-phosphate biosynthetic pathway.* Journal of the American Chemical Society, 2005. **127**(11): p. 3682-3683.
68. Raschle, T., N. Amrhein, and T.B. Fitzpatrick, *On the two components of pyridoxal 5'-phosphate synthase from Bacillus subtilis.* Journal of Biological Chemistry, 2005. **280**(37): p. 32291-32300.
69. Guédez, G., et al., *Assembly of the eukaryotic PLP-synthase complex from Plasmodium and activation of the Pdx1 enzyme.* Structure, 2012. **20**(1): p. 172-184.
70. Raschle, T., et al., *Reaction Mechanism of Pyridoxal 5'-Phosphate Synthase DETECTION OF AN ENZYME-BOUND CHROMOPHORIC INTERMEDIATE.* Journal of Biological Chemistry, 2007. **282**(9): p. 6098-6105.

71. Hill, R.E., et al., *The Biogenetic Anatomy of Vitamin B6 A 13C NMR Investigation of the Biosynthesis of Pyridoxol in Escherichia Coli.* Journal of Biological Chemistry, 1996. **271**(48): p. 30426-30435.
72. Yang, Y., et al., *Identification and function of the pdxY gene, which encodes a novel pyridoxal kinase involved in the salvage pathway of pyridoxal 5'-phosphate biosynthesis in Escherichia coli K-12.* Journal of Bacteriology, 1998. **180**(7): p. 1814-1821.
73. Yang, Y., G. Zhao, and M.E. Winkler, *Identification of the pdxK gene that encodes pyridoxine (vitamin B6) kinase in Escherichia coli K-12.* FEMS Microbiology Letters, 1996. **141**(1): p. 89-95.
74. Mooney, S., et al., *Vitamin B6: a long known compound of surprising complexity.* Molecules, 2009. **14**(1): p. 329-351.
75. Cao, P., et al., *Crystal structure of human pyridoxal kinase.* Journal of Structural Biology, 2006. **154**(3): p. 327-332.
76. Safo, M.K., et al., *Crystal structure of pyridoxal kinase from the Escherichia coli pdxK gene: implications for the classification of pyridoxal kinases.* Journal of Bacteriology, 2006. **188**(12): p. 4542-4552.
77. Safo, M.K., et al., *Crystal structure of the PdxY Protein from Escherichia coli.* Journal of Bacteriology, 2004. **186**(23): p. 8074-8082.
78. Musayev, F.N., et al., *Crystal structure of human pyridoxal kinase: Structural basis of M+ and M2+ activation.* Protein Science, 2007. **16**(10): p. 2184-2194.
79. Zhao, G. and M.E. Winkler, *4-Phospho-hydroxy-L-threonine is an obligatory intermediate in pyridoxal 5'-phosphate coenzyme biosynthesis in Escherichia coli K-12.* FEMS Microbiology Letters, 1996. **135**(2-3): p. 275-280.
80. Pédelacq, J.D., et al., *Crystal structure of a putative pyridoxine 5'-phosphate oxidase (Rv2607) from Mycobacterium tuberculosis.* PROTEINS: Structure, Function, and Bioinformatics, 2006. **62**(3): p. 563-569.
81. Biswal, B.K., et al., *The molecular structure of Rv2074, a probable pyridoxine 5'-phosphate oxidase from Mycobacterium tuberculosis, at 1.6 Å resolution.* Acta Crystallographica Section F: Structural Biology and Crystallization Communications, 2006. **62**(8): p. 735-742.
82. Sonenshein, A.L., *Control of key metabolic intersections in Bacillus subtilis.* Nature Reviews Microbiology, 2007. **5**(12): p. 917-927.

83. Chin, R.M., et al., *The metabolite α-ketoglutarate extends lifespan by inhibiting ATP synthase and TOR.* Nature, 2014. **510**(7505): p. 397-401.
84. Mouilleron, S. and B. Golinelli-Pimpaneau, *Conformational changes in ammonia-channeling glutamine amidotransferases.* Current Opinion in Structural Biology, 2007. **17**(6): p. 653-664.
85. Bauer, J.A., et al., *Three-dimensional structure of YaaE from Bacillus subtilis, a glutaminase implicated in pyridoxal-5′-phosphate biosynthesis.* Journal of Biological Chemistry, 2004. **279**(4): p. 2704-2711.
86. Douangamath, A., et al., *Structural evidence for ammonia tunneling across the (βα) 8 barrel of the imidazole glycerol phosphate synthase bienzyme complex.* Structure, 2002. **10**(2): p. 185-193.
87. Raschle, T., et al., *Intersubunit cross-talk in pyridoxal 5′-phosphate synthase, coordinated by the C terminus of the synthase subunit.* Journal of Biological Chemistry, 2009. **284**(12): p. 7706-7718.
88. Zalkin, H. and J.L. Smith, *Enzymes utilizing glutamine as an amide donor.* Advances in enzymology and related areas of molecular biology, 1998. **72**: p. 87-144.
89. Tesmer, J.J., et al., *The crystal structure of GMP synthetase reveals a novel catalytic triad and is a structural paradigm for two enzyme families.* Nature Structural Biology, 1996. **3**(1): p. 74-86.
90. Strohmeier, M., et al., *Structure of a bacterial pyridoxal 5′-phosphate synthase complex.* Proceedings of the National Academy of Sciences, 2006. **103**(51): p. 19284-19289.
91. Zhang, X., et al., *Structural insights into the catalytic mechanism of the yeast pyridoxal 5-phosphate synthase Snz1.* Biochemical Journal, 2010. **432**(3): p. 445-454.
92. Zein, F., et al., *Structural insights into the mechanism of the PLP synthase holoenzyme from Thermotoga maritima.* Biochemistry, 2006. **45**(49): p. 14609-14620.
93. Smith, A.M., et al., *Crystal structures capture three states in the catalytic cycle of a pyridoxal phosphate (PLP) synthase.* Journal of Biological Chemistry, 2015. **290**(9): p. 5226-5239.

94. Flicker, K., et al., *Structural and thermodynamic insights into the assembly of the heteromeric pyridoxal phosphate synthase from Plasmodium falciparum.* Journal of Molecular Biology, 2007. **374**(3): p. 732-748.
95. Neuwirth, M., et al., *Thermodynamic Characterization of the Protein− Protein Interaction in the Heteromeric Bacillus subtilis Pyridoxalphosphate Synthase.* Biochemistry, 2007. **46**(17): p. 5131-5139.
96. Ullah, N., et al., *Solution Structures and Dynamic Assembly of the 24-Meric Plasmodial Pdx1–Pdx2 Complex.* International Journal of Molecular Sciences, 2020. **21**(17): p. 5971.
97. Zhu, J., et al., *A new arrangement of (β/α) 8 barrels in the synthase subunit of PLP synthase.* Journal of Biological Chemistry, 2005. **280**(30): p. 27914-27923.
98. Kim, S. and K.-J. Kim, *Crystal structure of Mycobacterium tuberculosis Rv2606c: A pyridoxal biosynthesis lyase.* Biochemical and Biophysical Research Communications, 2013. **435**(2): p. 255-259.
99. Neuwirth, M., et al., *X-ray crystal structure of Saccharomyces cerevisiae Pdx1 provides insights into the oligomeric nature of PLP synthases.* FEBS letters, 2009. **583**(13): p. 2179-2186.
100. Matsuura, A., et al., *Crystal structure of pyridoxal biosynthesis lyase PdxS from Pyrococcus horikoshii.* Molecules and Cells, 2012. **34**(4): p. 407-412.
101. Robinson, G.C., et al., *Structural definition of the lysine swing in Arabidopsis thaliana PDX1: Intermediate channeling facilitating vitamin B6 biosynthesis.* Proceedings of the National Academy of Sciences, 2016. **113**(40): p. E5821-E5829.
102. Pettersen, E.F., et al., *UCSF Chimera—a visualization system for exploratory research and analysis.* Journal of Computational Chemistry, 2004. **25**(13): p. 1605-1612.
103. Gengenbacher, M., et al., *Vitamin B6 Biosynthesis by the Malaria Parasite Plasmodium falciparum BIOCHEMICAL AND STRUCTURAL INSIGHTS.* Journal of Biological Chemistry, 2006. **281**(6): p. 3633-3641.
104. Laskowski, R.A., et al., *PDBsum: Structural summaries of PDB entries.* Protein Science, 2018. **27**(1): p. 129-134.
105. Robert, X. and P. Gouet, *Deciphering key features in protein structures with the new ENDscript server.* Nucleic Acids Research, 2014. **42**(W1): p. W320-W324.

106. Roy, A., A. Kucukural, and Y. Zhang, *I-TASSER: a unified platform for automated protein structure and function prediction.* Nature Protocols, 2010. **5**(4): p. 725-738.
107. Müller, I.B., J.E. Hyde, and C. Wrenger, *Vitamin B metabolism in Plasmodium falciparum as a source of drug targets.* Trends in Parasitology, 2010. **26**(1): p. 35-43.
108. Wrenger, C., et al., *Analysis of the vitamin B6 biosynthesis pathway in the human malaria parasite Plasmodium falciparum.* Journal of Biological Chemistry, 2005. **280**(7): p. 5242-5248.
109. Fonda, M.L. and C. Harker, *Metabolism of pyridoxine and protein binding of the metabolites in human erythrocytes.* The American Journal of Clinical Nutrition, 1982. **35**(6): p. 1391-1399.
110. Loria, P., et al., *Inhibition of the peroxidative degradation of haem as the basis of action of chloroquine and other quinoline antimalarials.* Biochemical Journal, 1999. **339**(2): p. 363-370.
111. Blanchet, C.E., et al., *Versatile sample environments and automation for biological solution X-ray scattering experiments at the P12 beamline (PETRA III, DESY).* Journal of Applied Crystallography, 2015. **48**(2): p. 431-443.
112. Kikhney, A.G. and D.I. Svergun, *A practical guide to small angle X-ray scattering (SAXS) of flexible and intrinsically disordered proteins.* FEBS Letters, 2015. **589**(19): p. 2570-2577.
113. Mertens, H.D. and D.I. Svergun, *Structural characterization of proteins and complexes using small-angle X-ray solution scattering.* Journal of Structural Biology, 2010. **172**(1): p. 128-141.
114. Lipfert, J. and S. Doniach, *Small-angle X-ray scattering from RNA, proteins, and protein complexes.* Annu. Rev. Biophys. Biomol. Struct., 2007. **36**: p. 307-327.
115. Burkhardt, A., et al., *Status of the crystallography beamlines at PETRA III.* The European Physical Journal Plus, 2016. **131**(3): p. 56.
116. Grimes, J.M., et al., *Where is crystallography going?* Acta Crystallographica Section D: Structural Biology, 2018. **74**(2): p. 152-166.
117. Rupp, B., *Biomolecular crystallography: principles, practice, and application to structural biology*. Garland Science, 2009.

118. Rhodes, G. and J.B. Cooper, *Crystallography Made Crystal Clear*. 2006.
119. Wagner, A. *DLS-CCP4 Data Collection and Structure Solution Workshop*. 2016; Available from: https://www.ccp4.ac.uk/schools/DLS-2016/.
120. Sparta, K.M., et al., *XDSAPP2. 0.* Journal of Applied Crystallography, 2016. **49**(3): p. 1085-1092.
121. Blow, D., *Outline of crystallography for biologists*. Oxford University Press on Demand, 2002.
122. Gasteiger, E., et al., *Protein identification and analysis tools on the ExPASy server*, in *The proteomics protocols handbook*. 2005, Springer. p. 571-607.
123. Laemmli, U.K., *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-685.
124. Hames, B.D., *Gel electrophoresis of proteins: a practical approach*. Vol. 197. OUP Oxford, 1998.
125. Greenfield, N.J., *Using circular dichroism spectra to estimate protein secondary structure.* Nature Protocols, 2006. **1**(6): p. 2876.
126. Kelly, S.M., T.J. Jess, and N.C. Price, *How to study proteins by circular dichroism.* Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics, 2005. **1751**(2): p. 119-139.
127. Yang, J.T., C.-S.C. Wu, and H.M. Martinez, *Calculation of protein conformation from circular dichroism.* Methods in enzymology, 1986. **130**: p. 208-269.
128. Dierks, K., et al., *Dynamic light scattering in protein crystallization droplets: adaptations for analysis and optimization of crystallization processes.* Crystal Growth and Design, 2008. **8**(5): p. 1628-1634.
129. Falke, S. and C. Betzel, *Dynamic Light Scattering (DLS)*, in *Radiation in Bioanalysis*. 2019, Springer. p. 173-193.
130. Jeffries, C.M., et al., *Preparing monodisperse macromolecular samples for successful biological small-angle X-ray and neutron-scattering experiments.* Nature protocols, 2016. **11**(11): p. 2122-2153.
131. Franke, D., et al., *ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions.* Journal of applied crystallography, 2017. **50**(4): p. 1212-1225.

132. Konarev, P.V., et al., *PRIMUS: a Windows PC-based system for small-angle scattering data analysis.* Journal of applied crystallography, 2003. **36**(5): p. 1277-1282.
133. Svergun, D., *Determination of the regularization parameter in indirect-transform methods using perceptual criteria.* Journal of Applied Crystallography, 1992. **25**(4): p. 495-503.
134. Fischer, H., et al., *Determination of the molecular weight of proteins in solution from a single small-angle X-ray scattering measurement on a relative scale.* Journal of Applied Crystallography, 2010. **43**(1): p. 101-109.
135. Burchard, W., M. Schmidt, and W. Stockmayer, *Information on polydispersity and branching from combined quasi-elastic and intergrated scattering.* Macromolecules, 1980. **13**(5): p. 1265-1272.
136. Rambo, R.P. and J.A. Tainer, *Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law.* Biopolymers, 2011. **95**(8): p. 559-571.
137. Svergun, D.I., *Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing.* Biophysical Journal, 1999. **76**(6): p. 2879-2886.
138. Svergun, D.I., M.V. Petoukhov, and M.H. Koch, *Determination of domain structure of proteins from X-ray solution scattering.* Biophysical Journal, 2001. **80**(6): p. 2946-2953.
139. Petoukhov, M.V. and D.I. Svergun, *Global rigid body modeling of macromolecular complexes against small-angle scattering data.* Biophysical Journal, 2005. **89**(2): p. 1237-1250.
140. Svergun, D., C. Barberato, and M.H. Koch, *CRYSOL–a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates.* Journal of Applied Crystallography, 1995. **28**(6): p. 768-773.
141. Blanchet, C.E. and D.I. Svergun, *Small-angle X-ray scattering on biological macromolecules and nanocomposites in solution.* Annual Review of Physical Chemistry, 2013. **64**.
142. Jacques, D.A. and J. Trewhella, *Small-angle scattering for structural biology—Expanding the frontier while avoiding the pitfalls.* Protein Science, 2010. **19**(4): p. 642-657.

143. Mathew, E., A. Mirza, and N. Menhart, *Liquid-chromatography-coupled SAXS for accurate sizing of aggregating proteins.* Journal of Synchrotron Radiation, 2004. **11**(4): p. 314-318.
144. Graewert, M.A., et al., *Automated pipeline for purification, biophysical and x-ray analysis of biomacromolecular solutions.* Scientific Reports, 2015. **5**(1): p. 1-8.
145. Filipe, V., A. Hawe, and W. Jiskoot, *Critical Evaluation of Nanoparticle Tracking Analysis (NTA) by NanoSight for the Measurement of Nanoparticles and Protein Aggregates.* Pharmaceutical Research, 2010. **27**(5).
146. Gallego-Urrea, J.A., J. Tuoriniemi, and M. Hassellöv, *Applications of particle-tracking analysis to the determination of size distributions and concentrations of nanoparticles in environmental, biological and food samples.* TrAC Trends in Analytical Chemistry, 2011. **30**(3): p. 473-483.
147. Sonn-Segev, A., et al., *Quantifying the heterogeneity of macromolecular machines by mass photometry.* Nature communications, 2020. **11**(1): p. 1-10.
148. Verschueren, H., *Interference reflection microscopy in cell biology: methodology and applications.* Journal of cell science, 1985. **75**(1): p. 279-301.
149. Gast, K., *Dynamic and static light scattering.* Instrumental Analysis of Intrinsically Disordered Proteins: Assessing Structure and Conformation, 2010: p. 477-524.
150. Young, G., et al., *Quantitative mass imaging of single biological macromolecules.* Science, 2018. **360**(6387): p. 423-427.
151. Lee, S.F. and D. Klenerman, *Weighing one protein with light.* Science, 2018. **360**(6387): p. 378-379.
152. Brenner, S. and R. Horne, *A negative staining method for high resolution electron microscopy of viruses.* Biochimica et biophysica acta, 1959. **34**: p. 103-110.
153. Cheng, Y. and T. Walz, *The advent of near-atomic resolution in single-particle electron microscopy.* Annual review of biochemistry, 2009. **78**: p. 723-742.
154. De Carlo, S. and J.R. Harris, *Negative staining and cryo-negative staining of macromolecules and viruses for TEM.* Micron, 2011. **42**(2): p. 117-131.
155. Jensen, G., *Cryo-EM Part A: sample preparation and data collection*. 2010: Academic Press.

156. Ohi, M., et al., *Negative staining and image classification—powerful tools in modern electron microscopy.* Biological procedures online, 2004. **6**(1): p. 23-34.
157. Adams, P.D., et al., *PHENIX: a comprehensive Python-based system for macromolecular structure solution.* Acta Crystallographica Section D: Biological Crystallography, 2010. **66**(2): p. 213-221.
158. McCoy, A.J., et al., *Phaser crystallographic software.* Journal of Applied Crystallography, 2007. **40**(4): p. 658-674.
159. Afonine, P.V., et al., *Towards automated crystallographic structure refinement with phenix. refine.* Acta Crystallographica Section D: Biological Crystallography, 2012. **68**(4): p. 352-367.
160. Emsley, P. and K. Cowtan, *Coot: model-building tools for molecular graphics.* Acta Crystallographica Section D: Biological Crystallography, 2004. **60**(12): p. 2126-2132.
161. Joosten, R.P., et al., *The PDB_REDO server for macromolecular structure model optimization.* IUCrJ, 2014. **1**(4): p. 213-220.
162. Gore, S., et al., *Validation of structures in the Protein Data Bank.* Structure, 2017. **25**(12): p. 1916-1927.
163. Micsonai, A., et al., *Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy.* Proceedings of the National Academy of Sciences, 2015. **112**(24): p. E3095-E3103.
164. Cole, D., et al., *Label-free single-molecule imaging with numerical-aperture-shaped interferometric scattering microscopy.* ACS Photonics, 2017. **4**(2): p. 211-216.
165. Soltermann, F., et al., *Quantifying Protein–Protein Interactions by Molecular Counting with Mass Photometry.* Angewandte Chemie, 2020.
166. 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, 2010. **584**(19): p. 4169-4174.
167. Glatter, O., *A new method for the evaluation of small-angle scattering data.* Journal of Applied Crystallography, 1977. **10**(5): p. 415-421.
168. Putnam, C.D., et al., *X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures,*

conformations and assemblies in solution. Quarterly Reviews of Biophysics, 2007. **40**(3): p. 191-285.

169. Mukherjee, A. and V.A. Hackley, *Separation and characterization of cellulose nanocrystals by multi-detector asymmetrical-flow field-flow fractionation.* Analyst, 2018. **143**(3): p. 731-740.
170. Waterhouse, A., et al., *SWISS-MODEL: homology modelling of protein structures and complexes.* Nucleic Acids Research, 2018. **46**(W1): p. W296-W303.
171. Altschul, S.F., et al., *Basic local alignment search tool.* Journal of Molecular Biology, 1990. **215**(3): p. 403-410.
172. DeLano, W.L., *Pymol: An open-source molecular graphics tool.* CCP4 Newsletter on Protein Crystallography, 2002. **40**(1): p. 82-92.
173. Piiadov, V., et al., *SAXSMoW 2.0: Online calculator of the molecular weight of proteins in dilute solution from experimental SAXS data measured on a relative scale.* Protein Science, 2019. **28**(2): p. 454-463.
174. Weichenberger, C.X. and B. Rupp, *Ten years of probabilistic estimates of biocrystal solvent content: new insights via nonparametric kernel density estimate.* Acta Crystallographica Section D: Biological Crystallography, 2014. **70**(6): p. 1579-1588.
175. Liao, S.-M., et al., *The multiple roles of histidine in protein interactions.* Chemistry Central Journal, 2013. **7**(1): p. 1-12.
176. Krissinel, E. and K. Henrick, *Inference of macromolecular assemblies from crystalline state.* Journal of Molecular Biology, 2007. **372**(3): p. 774-797.
177. Tonkin, M.L., et al., *Structural and functional divergence of the aldolase fold in Toxoplasma gondii.* Journal of Molecular Biology, 2015. **427**(4): p. 840-852.
178. Organization, W.H., *Guidelines for the treatment of malaria*. World Health Organization, 2015.
179. Knöckel, J., et al., *The apicomplexan parasite Toxoplasma gondii generates pyridoxal phosphate de novo.* Molecular and Biochemical Parasitology, 2007. **152**(1): p. 108.
180. Dick, T., et al., *Vitamin B6 biosynthesis is essential for survival and virulence of Mycobacterium tuberculosis.* Molecular Microbiology, 2010. **78**(4): p. 980-988.

181. Divo, A.A., et al., *Nutritional requirements of Plasmodium falciparum in culture. I. Exogenously supplied dialyzable components necessary for continuous growth.* The Journal of Protozoology, 1985. **32**(1): p. 59-64.
182. Müller, I.B., et al., *The assembly of the plasmodial PLP synthase complex follows a defined course.* PLoS One, 2008. **3**(3): p. e1815.
183. Ventura, S., *Sequence determinants of protein aggregation: tools to increase protein solubility.* Microbial Cell Factories, 2005. **4**(1): p. 1-8.
184. Fernandez-Escamilla, A.-M., et al., *Prediction of sequence-dependent and mutational effects on the aggregation of peptides and proteins.* Nature Biotechnology, 2004. **22**(10): p. 1302-1306.
185. Chaudhuri, B.N., et al., *Crystal structure of imidazole glycerol phosphate synthase: a tunnel through a (β/α) 8 barrel joins two active sites.* Structure, 2001. **9**(10): p. 987-997.

Appendix

A1. List of Figures

A2. List of Tables

A3. Chemicals and Hazards

A3.1. Chemicals

Compound	CAS-No.	GHS hazard	Hazard statements*	Precautionary statements**
Acetic acid	64-19-7	GHS02, GHS05	H226, H314	P280, P305+351+338, P303
Acrylamide 37%	79-06-1	GHS06, GHS08	H301, H312, H315, H317, H319, H332, H340, H350, H316f, H372	P201, P280, P301+310, P305+351+338, P308+313
Agarose	9012-36-6	-	-	-
Ampicillin	69-52-3	GHS08	H317, H334	P280, P261, P302+352, P342+311
APS	7727-54-0	GHS03,GHS07, GHS08	H272, H302, H315, H317, H319, H334, H335	P280, P305+351+338, P302+352, P304+340, P301+312
Bromophenol blue	115-39-9	-	-	-
$CaCl_2$	10043-52-4	GHS07	H319	P305+351+338
Calcium acetate	114460-21-8	-	-	-
Citric acid	77-92-9	GHS07	H319	P264, P280, P305 + P351 + P338, P337 + P313
Coomassie Brilliant Blue R250	6104-59-2	-	-	-
DTT	577517	GHS07	H302, H315, H319,H335	P302,352, P305+351+338
EDTA	60-00-4	GHS07	H319	P305+351+338
Ethanol	64-17-5	GHS02	H225, H319	P210, P240, P305+351+338, P403+233
Ethidium bromide	1239-45-8	GHS06, GHS08	H302, H330, H341	P260, P281, P284, P310
Glucose	50-99-7	-	-	-
Glycerol	56-81-5	-	-	-
Glycine	56-40-6	-	-	-
Guanidine hydrochloride	50-01-1	GHS07	H302 + H332, H315, H319	P261, P264, P301 + P312, P302 + P352, P304 + P340, P312, P305 + P351 + P338
Hepes	7365-45-9	-	-	-
Hydrochloric acid >37%	7647-01-0	GHS05, GHS07	H290, H314, H335	P234, P261, P271, P280, P303 + P361 + P353, P305 + P351 + P338
Imidazole	288-32-4	GHS05,GHS07, GHS08	H302, H314, H360D	P260, P280, P301 + 312, P303 + 361 + 353, P304 + 340 + 310, P305 + 351 + 338
Isopropanol	67-63-0	GHS02,GHS07	H225, H319, H336	P210, P233, P305+351+338
KCl	7447-40-7	-	-	-
LiCl	7447-41-8	GHS07	H302, H315, H319	P264, P270, P280, P301 + P312, P302 + P352, P305 + P351 + P338
Lithium sulfate	10377-48-7	GHS07	H302, H319	P280, P301 + P312 + P330, P305 + P351 + P338, P337 + P313
Lithium citrate	313222-91-2	GHS07	H319	P305 + P351 + P338
Compound	**CAS-No.**	**GHS hazard**	**Hazard**	**Precautionary statements**

			statements	
$MgCl_2$	7786-30-03	-	-	-
NaH_2PO_4	7558-80-7	-	-	-
NaOH	1310-73-2	GHS05	H314	P280, P310, P305+351+338
NaCl	7647-14-5	-	-	-
Trisodium citrate	6132-04-3	-	-	-
Paraffin	8002-74-2	-	-	-
PEG4000	25322-68-3	-	-	-
PEG6000	25322-68-3	-	-	-
PEG8000	25322-68-3	-	-	-
PMSF	329-98-6	GHS05,GHS06	H301, H314	P280, P310, P361, P305+351+338,
SDS	151-21-3	GHS02,GHS05, GHS07	H228, H302, H315, H318, H335, H412	P210, P273, P280, P301 + P312 + P330, P302 + P352, P305 + P351 + P338 + P310
TEMED	110-18-9	GHS02,GHS05, GHS07	H225, H302, H314, H332	P261, P280, P305+351+338
Tetracycline	60-54-8	GHS07	H302	P301 + P312 + P330
Tris	1185-53-1	GHS07	H315, H319, H335	P261, P305+351+338
Tryptone	91079-40-2	-	-	-
Yeast Extract	8013-01-2	-	-	-
Casein	9000-71-9	-	-	-
K_2HPO_4	16788-57-1	-	-	-
KH_2PO_4	7778-77-0	-	-	-
Carbenicillin	4800-94-6	GHS08	H317, H334	P261, P272, P280, P284, P302 + P352, P304 + P340 + P312
Chloramphenicol	56-75-7	GHS08	H351	P201, P308+313
Isopropyl-β-D Thiogalacto pyranosid (IPTG)	367-93-1	-	-	-
Bicine	150-25-4	GHS07	H315, H319, H336	P321
Ammonium Acetate	631-61-8	-	-	-
Ammonium chloride	12125-02-9	GHS07	H302 - H319	P264, P270, P280, P301 + P312, P305 + P351 + P338, P337 + P313
Ammonium sulfate	7783-20-2	-	-	-
Bis-Tris	6976-37-0	GHS07	H315, H319, H335	P261, P305+351+338
MES	4432-31-9	GHS07	H315, H319, H336	P261, P305+351+339
Sodium Acetate	6131-90-4	-	-	-
Sodium Citrate	68-04-2	-	-	-
Sodium malonate	141-95-7	-	-	-
NaI	7681-82-5	GHS07,GHS08, GHS09	H315, H319, H372, H400	P260, P264, P273, P302 + P352, P305 + P351 + P338, P314

Compound	CAS-No.	GHS hazard	Hazard statements	Precautionary statements
$NiSO_4$	10101-97-0	GHS07,GHS08, GHS09	H302, H332, H315, H317, H334, H341, H350i,H360D, H372, H410	P201, P273, P280, P302+352, P304+340, P308+313
Ammonium hydrogen carbonate	1066-33-7	GHS07	H302	P301+312, P330
Acetonitrile	75-05-8	GHS02,GHS07	H225, H332, H302, H312, H319	P210, P240, P302+352, P305+351+ 338, P403+233
Formic Acid	64-18-6	GHS02, GHS05, GHS06	H226, H302, H314, H331,	P210, P280, P303+361+353, P304+340+310, P305+351+338, P403+233
TCEP	51805-45-9	GHS05	H314	P260, P280, P301 + P330 + P331, P303 + P361 + P353 - P304 + P340 + P310, P305 + P351 + P338
MTG	96-27-5	GHS06	H302, H311, H315, H317	P280, P301 + 312 + 330, P302 + 352. P312
R5P	207671-46-3	-	-	-
G3P	591-57-1	-	-	-
Glutamine	56-85-9	-	-	-

* "2" for physical hazards
"3" for health hazards
"4" for environmental hazards

** "1" for general precautionary statements
"2" for prevention precautionary statements
"3" for response precautionary statements
"4" for storage precautionary statements
"5" for disposal precautionary statements

A3.2. Crystallization Screens

Crystallization Screen	Supplier	GHS hazard	Hazard statements	Precautionary Statements
$AmSO_4$-Suite	Qiagen	GHS02, GHS06, GHS08, GHS09	H225, H301, H330, H350, H340, H360FD, H371, H411	P101, P201, P273, P280, P309+311
ComPAS-Suite	Qiagen	GHS02, GHS06, GHS07, GHS08, GHS09	H225, H301, H302, H315, H319, H331, H332, H335, H340, H350, H360FD, H373, H411	P101, P201, P270, P273, P280, P305+351+338, P309+311, P313
JCSG-Suite	Molecular Dimensions	GHS02, GHS05, GHS06,GHS07, GHS08, GHS09	H225, H301, H312, H315, H318, H331, H335, H350, H411	P101, P201, P270, P273, P280, P305+351+338, P309+311, P313
MORPHEUS	Molecular Dimensions	GHS02, GHS06, GHS07, GHS08, GHS09	H225, H301, H302, H315, H319, H331, H332, H335, H340, H350, H360FD, 361d, H373, H411	P101, P201, P270, P273, P280, P305+351+338, P309+311, P313
PACT premier	Molecular Dimensions	GHS06	H301, H331, H412	P101, P270, P273, P280, P309+311
Structure	Molecular Dimensions	GHS02, GHS06, GHS07, GHS08, GHS09	H225, H301, H302, H315, H319, H331, H332, H335, H340, H350, H360FD, 361d, H373, H411	P101, P201, P270, P273, P280, P305+351+338, P309+311, P313
Stura Footprint & Macrosol	Molecular Dimensions	GHS02, GHS06, GHS07, GHS08, GHS09	H225, H301, H302, H315, H319, H331, H332, H335, H340, H350, H360FD, H373, H411	P101, P201, P270, P273, P280, P305+351+338, P309+311, P313

A3.3. GHS and Risk Symbols

Figure 56: GHS Pictograms: Summary of all GHS pictograms along with the respective nomenclature.

A3.4. Hazard Statements

H225	Highly flammable liquid and vapour.
H226	Flammable liquid and vapour.
H228	Flammable solid.
H272	May intensify fire; oxidiser.
H280	Contains gas under pressure; may explode if heated.
H290	May be corrosive to metals.
H300	Fatal if swallowed.
H301	Toxic if swallowed.
H302	Harmful if swallowed.
H302+H332	Harmful if swallowed or if inhaled.
H310	Fatal in contact with skin.
H311	Toxic in contact with skin.
H312	Harmful in contact with skin.
H314	Causes severe skin burns and eye damage.
H315	Causes skin irritation.
H316	Causes mild skin irritation
H317	May cause an allergic skin reaction.
H318	Harmful if swallowed or if inhaled.
H319	Causes serious eye irritation.
H330	Fatal if inhaled.
H331	Toxic if inhaled.
H332	Harmful if inhaled.
H334	May cause allergy or asthma symptoms or breathing difficulties if inhaled.
H335	May cause respiratory irritation.
H336	May cause drowsiness or dizziness.
H340	May cause genetic defects.
H341	Suspected of causing genetic defects.
H350	May cause cancer.
H360fd	May damage fertility or the unborn child.
H361f	Suspected of damaging fertility.
H372	Cause damage to organs through prolonged or repeated exposure.
H373	May cause damage to organs through prolonged or repeated
H400	exposure.
H410	Very toxic to aquatic life with long lasting effects.
H411	Toxic to aquatic life with long lasting effects.
H412	Harmful to aquatic life with long lasting effects.

A3.5. Precautionary Statements

P101	If medical advice is needed, have product container or label at hand.
P201	Obtain special instructions before use.
P210	Keep away from heat, hot surfaces, sparks, open flames and other ignition sources. No smoking.
P233	Keep container tightly closed.
P234	Heating may cause a fire.
P240	Ground and bond container and receiving equipment.
P260	Do not breathe dust/fume/gas/moist/vapours/spray.
P261	Avoid breathing dust/fume/gas/mist/vapours/spray.
P264	Wash thoroughly after handling.
P270	Do not eat, drink, or smoke when using this product.
P271	May be harmful if inhaled.
P273	Avoid release to the environment.
P280	Wear protective gloves/protective clothing/eye protection/face protection.
P281	Suspected of causing genetic defects.
P284	Fatal if inhaled.
P301+310	IF SWALLOWED: Immediately call a POISON CENTRE or doctor/physician.
P301+312	IF SWALLOWED: Call a POISON CENTRE or doctor if you feel unwell.
P301+312 +330	IF SWALLOWED: Call a POISON CENTRE or doctor/physician if you feel unwell. Rinse mouth.
P301+330 +331	IF SWALLOWED: Rinse mouth. Do NOT induce vomiting.
P302+352	IF ON SKIN: Wash with plenty of soap and water.
P303	Causes severe skin burns and eye damage.
P303+361 +353	IF ON SKIN (or hair): Remove/Take off immediately all contaminated clothing. Rinse skin with water or shower.
P304+340	IF INHALED: Remove person to fresh air and keep comfortable for breathing.
P304+340 +310	IF INHALED: Remove person to fresh air and keep comfortable for breathing. Immediately call a POISON CENTRE or doctor/physician.
P305+351 +338	IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing.
P305+351 +338	IF IN EYES: Rinse cautiously with water for several minutes. Remove contact lenses, if present and easy to do. Continue rinsing. Immediately call a POISON CENTRE or doctor/physician.
P308+310	IF exposed or concerned: Immediately call a POISON CENTRE or doctor.
P308+313	IF exposed or concerned: Get medical advice/attention.
P309+311	IF exposed or you feel unwell: Call a POISON CENTRE or doctor/ physician.
P310	Immediately call a POISON CENTRE or doctor.
P312	May be harmful if swallowed.
P321	Specific treatment (see respective MSDS).
P330	Rinse mouth.
P331	Do NOT induce vomiting.

P337+313	Causes eye irritation.
P342+311	If experiencing respiratory symptoms: Call a POISON CENTRE or doctor.
P370+378	In case of fire: Use dry sand, dry extinguishing powder or alcohol resistant foam for extinction.
P391	Collect spillage.
P403	Store in a well-ventilated place.
P403+233	Store in a well-ventilated place. Keep container tightly closed.
P403+235	Store in a well-ventilated place. Keep cool.
P405	Store locked up.
P501	Dispose of contents / container in accordance with local / regional / national / international regulations.

A4. Acknowledgment

My special gratitude to my supervisor Prof. Dr. Christian Betzel for providing me an interesting project and sharing knowledge with me about X-ray crystallography. I would also like to thank him for providing me space at the extremely brilliant X-ray light source at DESY and the equipped UHH laboratories. It was a very inspiring experience working with him with all his valuable suggestions during these years.

I am very grateful to the Higher Education Commission (HEC), Pakistan for providing me funding support, under the project "Faculty Development Program BZU Multan (100 PhD Scholarships) (Prime Minister's Directive)".

I would like to thank Dr. Markus Perbandt for all his help and constructive discussions in protein co-crystallization and solving the crystal structure data. I also appreciate Prof. Dr. Carsten Wrenger from USP, Brazil for this productive collaboration and ideas exchange during my PhD studies.

Thanks to Dr. Sven Falke for all his help and support in learning how to analyze and process SAXS data. Especial thanks to Dr. Celestin Mudogo Nzanzu for all his valuable suggestions in moving forward with my PhD studies. Wishing him all the best and successes in his career. Thanks to Petra Belda for all the technical support.

Big thanks to the team of P11 and especially to beamline scientist Dr. Sofiane Saouane. Also, thanks to the team of P12 beamline for their full support in SAXS data measuring and other people from EMBL for coordination during the synchrotron experiments.

Many thanks to Guillaume Pompidor and the team of beamline P13, who facilitate me in the data collection and transfer of the collected data.

Thanks to the CSSB team Dr. Carolin Seuring (Core Facility Head) and Cornelia Cazey (Staff Scientist) for their welcoming approach in the electron microscopy analyses.

I am very grateful to the Refeyn team for mass photometry measurements by MP (Mass Photometer) instrument.

I would like to thank Hina Andaleeb and Nadine Werner and my colleagues from AG Betzel for all their help and support. My biggest thanks to Konrad Siggelkow for his patience and kindness at the beginning of my doctorate.

Obviously, this work would not be accomplished without my family's love and support. I am very grateful to my wife for all the sacrifices she made and her continuous moral support in my hard times I faced in these four years.

To all whose presence was equally important, so this work could be accomplished.

Thank you all!

Selbstständigkeitserklärung

Hiermit erkläre ich an Eides statt, dass ich die an der Universität Hamburg zur Promotion eingereichte Dissertation mit dem Titel "Strukturelle und biophysikalische Analyse von plasmodialen und bakteriellen multimerischen Pdx1 - Pdx2 Proteine und Komplexe" im Institut für Biochemie des Fachbereichs Chemie der Universität Hamburg ohne sonstige nicht angeführte Hilfe durchgeführt und bei der Abfassung der Dissertation keine anderen als die dort aufgeführten Hilfsmittel benutzt habe.

Ferner versichere ich, dass ich bisher an keiner in-oder ausländischen Universität ein Gesuch um Zulassung zur Promotion eingereicht und daher weder diese noch eine andere Arbeit als Dissertation vorgelegt habe.

Hamburg, 15. March 2021.

Najeeb Ullah