# *In silico* characterization of pyrophosphate-fructose 6-phosphate 1-phosphotransferase reveals dual catalytic role in restoring central carbon metabolism in *Candidatus* Liberibacter asiaticus

Sandra Kathott Prakash[1], Ashwani Kumar Sharma[1,*], Jitin Singla[1,*]

[1]Department of Biosciences and Bioengineering, Indian Institute of Technology Roorkee, Roorkee-247667, India

*Correspondence addresses:

**Ashwani Kumar Sharma,** Department of Biosciences and Bioengineering, Indian Institute of Technology Roorkee, Roorkee-247667, India. Email: ashwani.sharma@bt.iitr.ac.in; Phone: +91-1332-285657; Fax: +91-1332-273560

**Jitin Singla,** Department of Biosciences and Bioengineering, Indian Institute of Technology Roorkee, Roorkee-247667, India. Email: jsingla@bt.iitr.ac.in; Phone: +91-1332-285609;

**Abstract:** *Candidatus* Liberibacter asiaticus (CLas) is a host-dependent obligate bacteria that causes citrus greening disease in citrus plants. The prolonged association of CLas with the host led to the loss of essential metabolic genes, including transaldolase in the pentose phosphate pathway, hindering the development of axenic culture in the host-free environment. In this study, *in silico* characterization of pyrophosphate-fructose 6-phosphate 1-phosphotransferase (PFP) protein was performed by utilizing sequence analysis, structure prediction, and molecular modeling studies, revealing its potential role in bypassing the loss of transaldolase. The findings suggest how PFP and fructose bisphosphate aldolase (FBPA) can reconstruct the disrupted central carbon metabolism and indirectly help CLas in ATP production. The molecular dynamics studies and free energy calculations of PFP-complexes showed binding affinity towards fructose 6-phosphate and sedoheptulose 7-phosphate, showing its dual function. The sequence-based analysis and metabolic modeling study also highlight the presence of other alternate enzymes or transport proteins in CLas, elucidating the broader metabolic resilience. This is the first report to provide structural insights into how PFP compensates for the lost metabolic function and survival strategies adopted by CLas despite its restrained metabolic capability.

**Keywords:** pentose phosphate pathway, transaldolase, pyrophosphate-fructose 6-phosphate 1-phosphotransferase, molecular modeling, metabolic resilience

## 1. Introduction

*Candidatus* Liberibacter asiaticus (CLas) is a gram-negative, phloem-limited alphaproteobacteria transmitted by the insect vector *Diaphorina citri*, commonly known as *psyllid*, causing the disease called "citrus greening" or Huanglongbing (HLB) in *citrus spp* (Bové, 2006). Major symptoms shown by HLB-infected citrus plants are leaf spots, malformed/discolored fruits, increased premature fruit abscission, and mortality rate (Bové, 2006). *Candidatus* Phytoplasma, a species homologous to CLas, is a phloem-limited, mycoplasma-like organism that lacks a cell wall (Montano et al., 2024). It has been reported to be associated with mixed infection with CLas and *Spiroplasma citri* to cause HLB in different citrus species (Noorizadeh et al., 2022). HLB affects citrus production worldwide, causing a decline in the economy in the regions of Asia, Africa, Oceania, and the Americas (Merfa et al., 2019).

Alphaproteobacteria is a diverse class of prokaryotes that have a significant role in the origin of mitochondria through a symbiotic relationship with the primitive eukaryotic cell (John & Whatley, 1975). Murray and Schleifer, two authors in 1995, proposed giving provisional status to the unculturable prokaryote taxa, and the International Committee on Systematic Bacteriology (ICSB) recommended assigning the term *Candidatus,* which is equivalent to genus (Murray & Stackebrandt, 1995). Various Liberibacter species belonging to the class alphaproteobacteria and classified under *Candidatus* provisional status include CLas, *Candidatus* Liberibacter africanus (CLaf), *Candidatus* Liberibacter americanus (CLam), *Candidatus* Liberibacter brunswickensis(CLbr), *Candidatus* Liberibacter europaeus (CLeu), *Candidatus* Liberibacter ctenarytainae (CLct), and *Candidatus* Liberibacter solanacearum (CLso). *Liberibacter crescens* (Lcr) strain BT-1 was the first culturable bacteria of the genus Liberibacter spp., isolated from Babaco Mountain papaya's infected leaf sap (*Carica stipulata* x *C. pubescens*) (Leonard et al., 2012). Lcr BT-1 is non-pathogenic and does not have any insect or plant host (Jain, Cai, et al., 2019a).

Several trials for the pure culture of CLas have been made, but either it failed or can't maintain the viability of the bacteria (Davis et al., 2008; Ha et al., 2019; Sechler et al., 2009). The essential genes that are crucial for the growth and survival of an organism can be affected by the changing environmental conditions and ecological adaptations, which may lead to either alteration or loss of these genes over time (Rancati et al., 2018). In case of CLas, the prolonged association with the host leads to genome minimization, hence losing many essential genes (Cai et al., 2022). CLas has 137 transport genes, of which 40 are ATP-binding cassette (ABC)

transporters capable of importing essential metabolites and thereby hampering the metabolic requirement of the citrus host, eventually leading to the death of the plant (N. Wang & Trivedi, 2013).

CLas has three types of amino acid transporter systems identified: amino acid ABC transporter permease, branched-chain amino acid transporter, and dicarboxylate/amino acid: cation symporter (Lin et al., 2011). A study from our lab had reported the alternate mechanism of putative cystine binding protein from CLas strain psy62, which showed a broad specific affinity towards amino acids other than cationic amino acids, and mutation studies were carried out to understand unique features of the protein (Kumar et al., 2019, 2020). Another study on the structural and biophysical characterization of a solute binding protein from CLas revealed the mechanism by which it transports two metal ions, $Mn^{2+}$ and $Zn^{2+}$ (Sharma et al., 2016). CLas have genes encoded to transport vitamins like thiamine ABC transporter permease ThiP, ThiQ, ThiB for transporting thiamine (Vitamin $B_1$) from the host (Lin et al., 2011). Genome analysis studies highlighted the absence of chorismate synthase, which produces precursor, chorismate, for the biosynthesis of folate (Vitamin $B_9$) and the absence of a transport system (Attaran et al., 2020). Additionally, CLas utilizes the secretory protein peroxiredoxin to reduce the hydrogen peroxide ($H_2O_2$) level (Bernardini et al., 2022; Jain, Munoz-Bodnar, et al., 2019b). In a previous study, our lab has characterized the bacterioferritin comigratory family protein, 1-cys peroxiredoxin, which can protect against peroxide-mediated cell damage (Singh et al., 2017). Even though individual proteins from CLas have been studied, several experimental approaches, like functional characterization, metabolomics, and genetic manipulation, are limited due to the inability to grow the bacteria in a host-free environment.

Previous studies have shown that CLas lacks essential genes associated with the glycolysis pathway, including glucose-6-phosphate isomerase (*pgi*), and this reaction can be compensated via the pentose phosphate pathway (PPP), but this possibility is compromised due to the absence of transaldolase (Attaran et al., 2020). The comparative genome analysis of CLso strain revealed the lack of phosphoenolpyruvate-phosphotransferase system (PTS), presence of partial PPP, and absence of energy metabolism related enzymes, including polyphosphate kinase, inorganic diphosphatase, terminal oxidases: cytochrome c oxidase and cytochrome bd complex (Lin et al., 2011). Additionally, the absence of gene encoding for methylglyoxal detoxification system suggests that the bacteria is incapable of producing its own adenosine triphosphate (ATP) (Jain et al., 2017). In 2010, Vahling *et al.* concluded that the ATP

requirement of the bacteria is fulfilled by importing the ATP from the host using translocase (GU011685.1) (Vahling et al., 2010). In this study, we discussed the possibility of CLas producing its own ATP by finding alternative metabolic reactions.

Flux distribution predictions using RNA sequencing data highlight the reliance of CLas on the host for maintaining an active PPP, underscoring its ecological and energy dependence on the host environment (Zuñiga et al., 2020). Hence, a functional PPP can be present in CLas under certain host-dependent constraints. The absence of transaldolase hampers the synthesis of erythrose 4-phosphate (E4P) and disrupts the later reaction to form glyceraldehyde 3-phosphate (G3P) and fructose 6-phosphate (F6P) from xylulose 5-phosphate (Xu5P) (Stincone et al., 2015). The multiple knockout study involving *E coli* transaldolase mutants revealed the dependence of bacteria on ATP-dependent phosphofructokinase (PFK-1) and fructose bisphosphate aldolase (FBPA) to complete PPP (Nakahigashi et al., 2009). The metabolome analysis also revealed the presence of metabolites, sedoheptulose 1,7-bisphosphate (SBP) and sedoheptulose 7-phosphate (S7P) (Nakahigashi et al., 2009). Hence, it is possible that CLas, which lacks a gene encoding for transaldolase, utilizes a homologous protein to compensate for the loss.

The present study focuses on an intermediate reaction involving pyrophosphate-fructose 6-phosphate 1-phosphotransferase (PFP) protein (homologous to PFK-1) that can complete the disrupted central carbon metabolism and compensate for the loss of the gene encoding for transaldolase. To annotate the biological function of the PFP protein, *in silico* structure-based studies, including structure prediction, molecular docking, and simulations of PFP in complex with substrates, are performed. The free energy calculation and per-residue decomposition analysis were employed to calculate the free energy of binding and to predict the residues that form a stable binding. The significance of pyrophosphate (PPi) and magnesium ion ($Mg^{2+}$) for the functional activity of PFP protein is also studied using molecular docking and simulations. We assessed the possibility of CLas producing its own ATP by finding alternative metabolic reactions, survival strategies, and the influence of ecological adaptations. Furthermore, sequence-based analysis is employed to find the alternative reactions that could compensate for the lost metabolic function, which have not been previously reported.

## 2. Results and Discussion

### 2.1. Sequence analysis

The sequence similarity search revealed that CLas only has a single enzyme, pyrophosphate-fructose 6-phosphate 1-phosphotransferase (PFP) (EC 2.7.1.90), which plays an essential role in the glycolytic pathway and belongs to the phosphofructokinase type A (IPR050929) known as PFKA. Metabolic model building for CLas bacteria using Kbase software (Arkin et al., 2018) with minimal glucose media showed the presence of the SBP reaction. Microorganisms lacking transaldolase (EC 2.2.1.2) may rely on the SBP pathway due to the reaction's reversibility compared to ATP-dependent phosphofructokinase (PFK-1) (EC 2.7.1.11) and also help in conserving energy substrates in an ATP-limited environment (Koendjbiharie et al., 2020). The presence of Npt1/Npt2 family nucleotide transporter enzyme, encoded for ATP:ADP translocase activity, implies the dependence of CLas on the host for energy needs. This suggests an evolutionary route adopted by CLas to regulate the utilization of ATP.

The bioinformatics tool ProtParam inferred stable protein with theoretical pI, extinction coefficient, and molecular weight of 6.56, 34755$M^{-1}$ $cm^{-1}$, and 44.4kDa, respectively. The sequences considered in the study are of *Liberibacter spp., Phytoplasma spp.,* and other species reported to encode PFP protein involved in the alternate reaction, i.e., Sequence similarity search showed *Clostridium spp.,* and *Entamoeba histolytica* encode for two or three types of PFK with a variance in the motif, showing specificity towards different phosphoryl donors (PPi, ATP, ATP-dependent variant with atypical amino acid combination denoted as PFK*). Multiple sequence alignment of PFP sequence from CLas shows variable identity with *Liberibacter spp.,* viz., CLaf (85.2%), CLbr (83.4%), CLso (81.5%), CLeu (76.1%), CLct (75.4%), CLam (74.5%), and Lcr BT-1 (64.8%). On comparing, *Phytoplasma spp.,* which possess ATP-dependent PFK, showed lower identity, viz., *Ca*.P. asteris (30.4%), *Ca*.P. australiense (30.3%), *Ca*.P. phoenicium (29.6%), *Ca*.P. luffae (27.9%), *Ca*.P. solani (27.2%), and *Ca*.P. melaleuceae (26.9). Other homologs included *E. coli* (27.9%), *Prevotella copri* (28.7%), *E.histolytica* (PFP: 24%, PFK*: 28%), *Clostridium thermosuccinogenes* (PFP: 27%, PFK-1: 24% PFK*: 26%), *Clostridium thermocellus* (PFP: 26%, PFK-1: 24%).

Besides the variable identity among the sequences, the multiple sequence alignment infers the conservation of the catalytic domain with less functional change for the binding of the substrate (**Figure 1**). The synonymous and non-synonymous substitutions of residues involved in the binding of ATP/PPi were observed in *Phytoplasma spp*., *P.copri, E.coli*, *C.thermosuccinogenes, C.thermocellus,* and *E.histolytica.* The *E.coli* PFK-1 protein has two

motifs, GGD**G** (101-104) and **G**TIDND (124-129), conserving the residue glycine at positions 104 and 124, which plays a major role in the active binding of ATP (Chi & Kemp, 2000). The sequence analysis across PFP proteins revealed the missense mutation of these sites with the signature patterns, GG(D/N)**D** and **K**TIDXD. In CLas PFP protein, corresponding residues were D123 and K149, making it dependent on PPi. The mutation studies revealed PFP from *E.histolytica* showed the activity to interact with ATP on double mutation of the corresponding residues from the motif to glycine (D175→G and K201→G) and thereby lost activity with PPi (Chi & Kemp, 2000). This suggests the loss or gain of polar contacts and charges, which influence the binding of ATP or PPi in different homologous proteins.

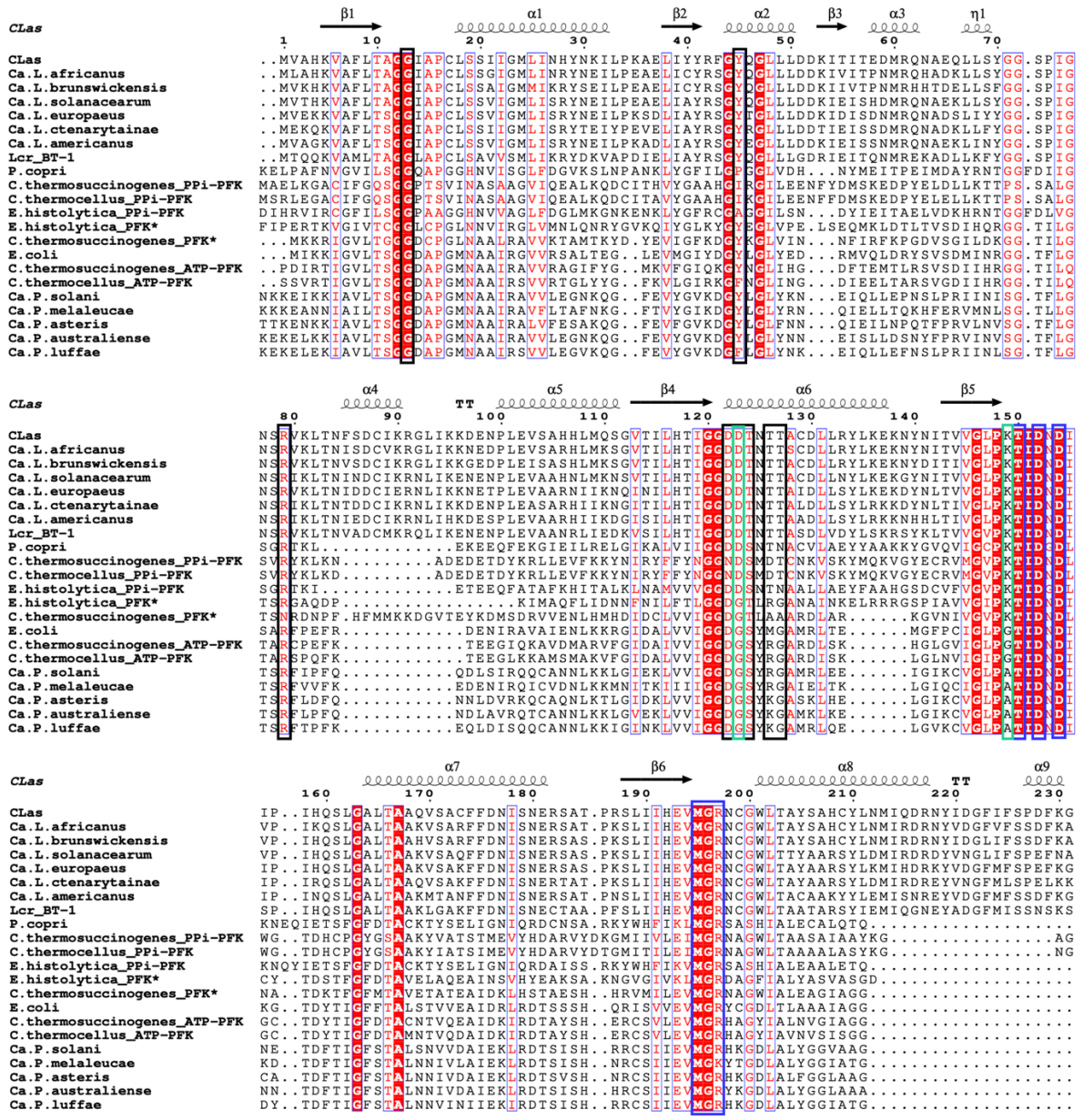

**Figure 1.** Multiple sequence alignment of PFK protein sequences from homologous species. The active sites are highlighted in black and blue boxes. The substrate binding sites are shown in blue boxes, and the PPi binding sites are indicated in black boxes. The green box highlights the atypical amino acid combination observed among different species.

Additionally, another variant of PFK, showing affinity towards ATP with motifs GGD**G** and **K**TIDXD, was observed in *C. thermosuccinogenes* and *E. histolytica.* Similar patterns were previously reported in kinetoplasts, *Leishmania donovani* and *Trypanosoma brucei* belonging to the eukarya domain (López et al., 2002). Our sequence analysis also found a non-canonical amino acid pattern of GGD**G** and **A**TIDXD motifs in PFK-1 of *Phytoplasma spp*., which showed glycine is substituted with alanine, which varies from the well-conserved PFK-1 and PFK* motifs. In *E. coli* PFK-1, glycine at position 108 (equivalent to T127 in CLas) is essential for accommodating the adenine moiety of ATP in the pocket (Moore et al., 2002a). The sequence analysis showed that glycine residue in PFK-1 and PFK* were substituted with aspartate or asparagine in PFP. Comparing with CLas, other substitutions in the active site were observed, including Y45→P/I/A/F, R79→N, T126→D/R/K/A/M, R197→K. The mutation of the active site residues of PFK during evolution has modulated the specificity towards different phosphoryl donors.

### 2.2. Phylogenetic tree and Molecular clock hypothesis

To understand the evolutionary divergence pattern and infer ancestral points, a phylogenetic tree was constructed using the maximum likelihood method in MEGA11 software (Tamura et al., 2011). The study utilized 22 PFK protein sequences, and the obtained tree possesses a maximum likelihood of -8935.3 and a genetic distance of 0.5. A mid-point rooted tree showed two major ancestral nodes from which PFK specific towards PPi and ATP were diverged (**Figure 2**). The CLas PFP is closely related to CLbr and CLaf, with the latter forming a sister clade. Comprehensive genome analysis reveals Lcr BT-1 forms a basal lineage with *Liberibacter spp.,* diverging as an outgroup (Batarseh et al., 2023). The tree depicts that PFP from *Liberibacter spp.,* and *Clostridium spp.,* are diverged from the same last common ancestor (LCA). Also, PFP from *E.histolytica* and *P.copri* are closely related, and this clade diverged from the same LCA as the homologs with PFK* and PFK-1. This infers the possibility of substitutions in the active site domain leading to a change in the affinity towards ATP/PPi. A previous study has suggested the occurrence of horizontal gene transfer events and has

analyzed the mutations seen in the active site that influenced the activity of PFK towards different phosphoryl donors (Bapteste et al., 2003). The tree showed three lineages descending (polytomy) from the same node, including PFK-1 from *Clostridium spp., E.coli, and Phytoplasma spp*., and this clade is diverged from the same LCA as *C.thermosuccinogenes* PFK*. Thus, *C.thermosuccinogenes* PFK* is a sister clade to the rest of the ingroup of PFK-1 from homologs, suggesting the early divergence of PFK*. Previous studies discussing PFK* or PFK-1 as the ancestral protein, considering the reversibility of reaction, function, mutations, and adaptive evolution among homologs, give contradicting opinions (Bapteste et al., 2003; Siebers et al., 1998). The current study has included the PFK protein sequences from homologous species lacking transaldolase and surviving in limited-oxygen or anerobic conditions except for *E.coli* (facultative anerobe) and Lcr BT-1 (aerobe). The results navigate the presence of different PFK variants in different species, including cases where two or three variants are encoded in the genome of the same species (*C.thermosuccinogenes* and *E.histolytica*), indicating the complex evolutionary pattern of PFK. Even though *Liberibacter spp*., and *Phytoplasma spp*., are homologous and cause mixed HLB infections, both encode for varying PFK and show divergence at different evolutionary times. Further analysis revealed a sister clade of *E.histolytica* PFP and *P.copri* PFP showing early divergence from the rest of the homologs encoding PFK* and PFK-1 from the same LCA. Hence, it is plausible that PFP protein is ancestral. The studies involving single mutation (D175→G) and double mutation (D175→G and K201→G) of residues from *E.histolytica* PFP led to gaining affinity towards ATP, which suggests ATP binding activity is encoded inherently in the gene (Chi & Kemp, 2000). Life started in an anoxic environment, and living forms evolved by taking protective strategies against oxygen tolerance when cyanobacteria produced oxygen by oxygenic photosynthesis ~2.4 billion years ago (Khademian & Imlay, 2021; Martin & Sousa, 2015). Notably, PFP is mainly present in anaerobic or microaerophilic, ATP-limited, intracellular prokaryotes, as well as in eukaryotes. Moreover, unlike PFK-1, PFP catalyzes a reversible reaction, lacks an allosteric site of regulation, and plays a multifunctional role as part of the SBP pathway, supporting the possibility that PFP could be ancestral (Compton & Patrick, 2025). The current data is still lacking to conclude that PFP is ancestral, but our study suggests the possibility.

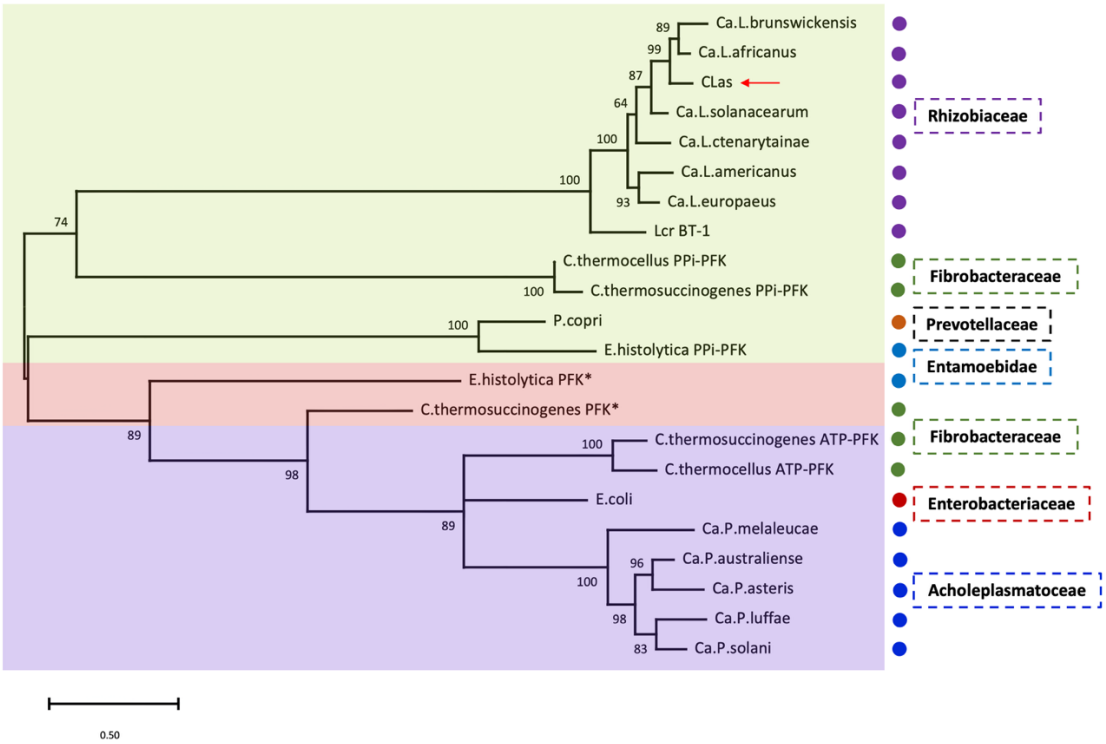


**Figure 2.** The phylogenetic tree was constructed using the maximum likelihood method in MEGA11 software with 1000 bootstraps. CLas PFP protein is highlighted with a red arrow, and the diverged lineages are categorized based on the family taxonomy level and represented with colored dots. The PFK variants from different homologs including PPi-dependent PFP (light green), ATP-dependent PFK* (light red), which has a binding motif similar to PFP, and ATP-dependent PFK-1 (purple), are highlighted in different colored boxes.

The Tajima's relative rate test was conducted to test the molecular clock hypothesis or null hypothesis on PFK sequences. CLas PFP protein was considered as the query taxa (taxa 1), other homologs (taxa 2) as the taxa to which molecular differences are measured, and *E.histolytica* PFK* as the outgroup. The Tajima's relative rate is a non-parametric calculation using the chi-square ($\chi^2$) approximation method with one degree of freedom, measuring the unique mutations in taxa 1 and taxa 2, denoted as $m_1$ and $m_2$ (Tajima, 1993). The null hypothesis is rejected if P ($\chi^2 \geq 3.841$) = 0.05 at 5% level. The more closely related taxa to CLas were CLaf and CLbr, with 2 and 3 unique mutations, respectively. CLct and Lcr BT-1 taxa showed the P-value of 1 with equal number of unique mutations. The test results showed that the evolution rate is constant for every ortholog except *C.thermosuccinogenes* PFK*, with a $\chi^2$ value of 9.78 and a P-value of 0.0017 (**Table 1**). Thus, CLas PFP and *C.thermosuccinogenes* PFK* evolved at different rates, with the highest number of substitutions in CLas PFP (m1 =

61). The inconsistent divergence rate also suggests PFK* has less conserved active site, which could be a transition towards PFK-1. Also, *C.thermocellum* PFK-1 appeared to have the highest unique mutation of 53, $\chi^2$=3.68, and P-value of 0.05 after *C.thermosuccinogenes* PFK*. The null hypothesis test revealed that PFK* exhibits high divergence compared to PFP and PFK-1.

| Name of the species | $\chi^2$ | P-value | $m_1$ | $m_2$ |
|---|---|---|---|---|
| *Ca.*L. africanus | 1.29 | 0.25 | 2 | 5 |
| *Ca.*L. americanus | 0.22 | 0.63 | 10 | 8 |
| *Ca.*L. brunswickensis | 0.50 | 0.47 | 3 | 5 |
| *Ca.*L. europaeus | 0.60 | 0.43 | 6 | 9 |
| *Ca.*L. ctenarytainae | 0 | 1 | 9 | 9 |
| *Ca.*L. solanacearum | 1.67 | 0.19 | 10 | 5 |
| *Liberibacter crescens* | 0 | 1 | 11 | 11 |
| *Ca.*P. asteris | 0.56 | 0.45 | 35 | 29 |
| *Ca.*P. australiense | 0.78 | 0.37 | 35 | 28 |
| *Ca.*P. luffae | 1.39 | 0.23 | 41 | 31 |
| *Ca.*P. solani | 1.92 | 0.16 | 37 | 26 |
| *Ca.*P. melaleuceae | 2.1 | 0.14 | 47 | 34 |
| *E. coli* | 2.8 | 0.09 | 42 | 28 |
| *Prevotella copri* | 0.80 | 0.37 | 44 | 36 |
| *Entamoeba histolytica*_ PFP | 0.05 | 0.83 | 45 | 43 |
| *Clostridium thermosuccinogenes*_PFP | 1.95 | 0.16 | 31 | 43 |
| *Clostridium thermosuccinogenes*_PFK-1 | 2.71 | 0.09 | 49 | 34 |
| ***Clostridium thermosuccinogenes*_PFK*** | **9.78** | **0.0017** | **61** | **31** |
| *Clostridium thermocellum*_PFP | 0.64 | 0.42 | 35 | 42 |
| *Clostridium thermocellum*_PFK-1 | 3.68 | 0.05 | 53 | 35 |

**Table 1.** The molecular clock hypothesis test showing divergence of each ortholog species (taxa 2) against CLas PFP protein (taxa 1) measured along with *E.histolytica* PFK* as outgroup. The non-parametric approximation calculated using the equation: $\chi^2 = (m_1-m_2)^2/(m_1+m_2)$, in which $m_1$ and $m_2$ denote unique differences found in taxa 1 and taxa 2.

### 2.3. Molecular modeling of CLas PFP and effective free energy calculation

The sequence similarity search against PDB database found PFK-1 from *E. coli* (PDB ID: 1PFK), which showed higher identity (28%), which is used as a reference structure for structural analysis. The sequence alignment of CLas PFP protein and structurally characterized PFK homologs revealed the conserved active site except for the ATP/PPi binding motif (**Supplementary Figure 1**). The protein structure of PFP was predicted using AlphaFold3 (in which PFP bound to $Mg^{2+}$ was predicted) (Abramson et al., 2024), ESMFold (Fang et al.,

2023), and RoseTTAFold (Krishna et al., 2024), and a comparison of predicted structures revealed the overlap in the catalytic domain shown in **Figure 3A**. The RMSD values calculated for the aligned structures on comparing with the catalytic site of PFK-1 from *E. coli* (PDB ID: 1PFK) were 1.46 A°, 1.56 A°, and 1.5 A° for AlphaFold3, ESMFold, and RoseTTAFold, respectively. For further study, the structure predicted using Alphafold3 was used.

As discussed in section **2.1**, sequence analysis of PFK had shown a distinct ATP/PPi binding domain. The crystal structure 1PFK from *E. coli* showed the conformational change occurs on the phosphoryl transfer to the F6P and binds with $Mg^{2+}$ in the active site (Shirakihara & Evans, 1988). To understand the structural aspects of the ATP/PPi binding motif, the predicted CLas PFP model was compared with 1PFK structure bound with ADP and $Mg^{2+}$. The AlphaFold3 model comprises α/β domain forming fold to accommodate substrate and inorganic phosphate (PPi) and $Mg^{2+}$. The active site residues predicted for PFP protein using NCBI conserved domain search tool were G12, G13, Y45, R79, D122, D123, T124, T126, T127, T150, D152, D154, M195, G196, R197, E267, Y324, and R327. The previous studies of *E. coli* PFK-1 revealed that the residues involved in the binding of the nucleotide moiety in the active pocket include Y41, F76, R77, R82, and R111. (X. Wang & Kemp, 1999). To coordinate the binding of the adenine moiety of ATP in *E. coli* PFK-1, a motif consisting of residues FPEFR (73-77) is required, which is missing in CLas PFP (Moore et al., 2002b). The secondary structure element depicted for FPEFR motif of 1PFK structure was a $3_{10}$ helix (Shirakihara & Evans, 1988), which is changed to a segment of loop in the active site of CLas PFP, showing the alteration of secondary structure in the influence of ATP/PPi (**Figure 3B**). The 1PFK was compared with the PFK protein from *Borrelia burgdorferi* (PDB ID:1KZH), and the predicted model of *Prevotella copri* and CLas PFP, which shows the distinct FPEFR motif of 1PFK (**Supplementary Figure 2**). The RMSD values showed for the superimposed structures of CLas PFP model when comparing with the 1PFK, 1KZH, and *P.copri* PFP model were 1.46 A°, 4.34 A° and 2.65 A°, respectively. Another key amino acid involved in the binding of ATP is G108 in 1PFK, which is substituted with aspartate (T127) in CLas PFP, influencing the hydrophobicity of the pocket (Moore et al., 2002b). Further, AlphaFold3 model was docked with PPi (PFP+$Mg^{2+}$) and performed molecular simulation for 300ns. The simulated model of PFP+$Mg^{2+}$ showed PPi fitting in the catalytic pocket, forming polar interactions with the residues S78, V80, D123, and T124, and $Mg^{2+}$ coordinated with D122, T150, D152, and hydrogen bonded with D154 (**Figure 3C**). It is reported that in *E. coli* PFK-1, $Mg^{2+}$ is primarily

coordinated by D103 and D129 (Shirakihara & Evans, 1988), which corresponds to D122 and D154 in CLas PFP. To assess the role of $Mg^{2+}$, simulation of CLas PFP structures bound to substrate in the absence and presence of $Mg^{2+}$ is conducted.

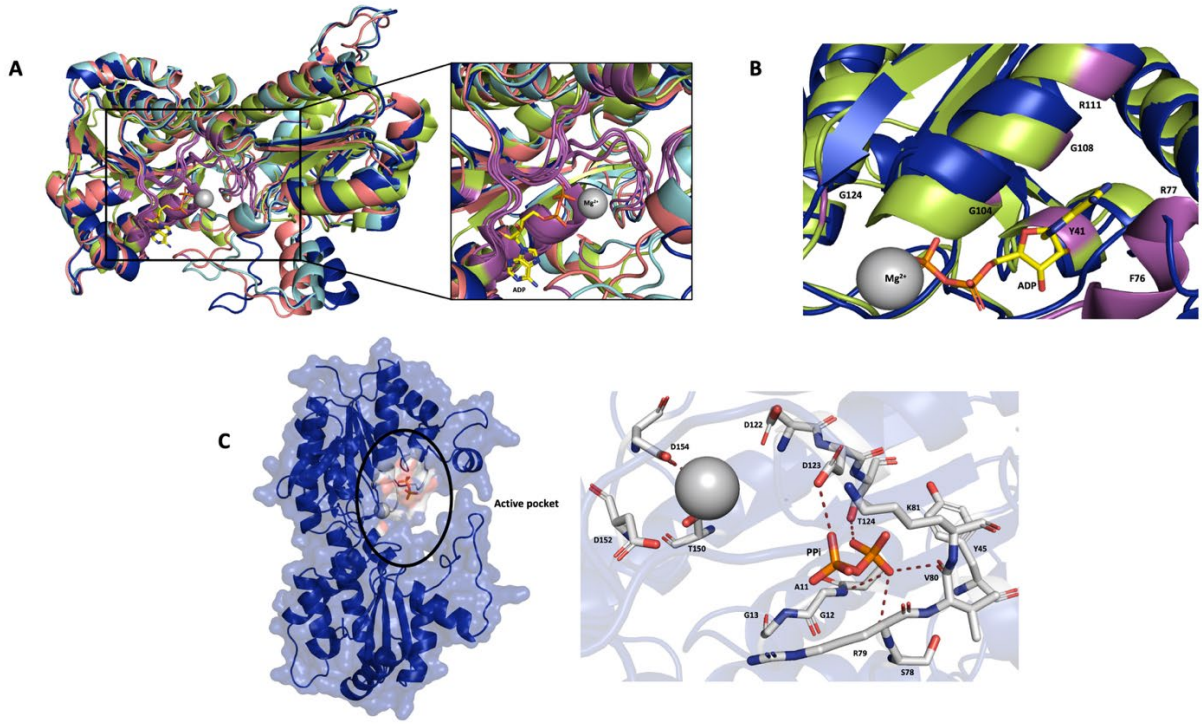


**Figure 3.** The PFP structure highlights the α/β domain and the catalytic domain found with PPi and $Mg^{2+}$ **(A)** The superimposed structures of PFP predicted using AlphaFold3 (blue), ESMFold (cyan), RoseTTAFold (pink), and *E. coli* PFK-1 (1PFK, green) are shown in cartoon representation, and the catalytic domain is highlighted in magenta color. (**B**) structural comparison of 1PFK with CLas PFP model highlighting the nucleotide binding site (magenta color) **(C)** The predicted model of the CLas PFP is shown in a cartoon representation in blue color, along with the catalytic domain shown on the right side. The active site residues highlighting the PPi and $Mg^{2+}$ bound region are shown on the right side. The hydrogen bond interactions are shown in red dotted lines.

Previous studies have reported that PFP can circumvent the loss of transaldolase by helping in the conversion of S7P to SBP and later forming erythrose 4-phosphate (E4P), catalyzed by fructose bisphosphate aldolase (FBPA) (Garschagen et al., 2021),(Koendjbiharie et al., 2020),(Susskind et al., 1982). Molecular modeling of PFP involving S7P as substrate is conducted to assess the function of PFP in catalyzing the unusual intermediate reaction to form SBP. The substrates used in the study were F6P (known substrate) (Shirakihara & Evans, 1988;Kricke et al., 1999) and S7P. The rigid docking revealed that F6P and S7P formed a

complex with the CLas PFP structure with a binding energy of -4.8 and -4.9 kcal/mol, and with CLas PFP+$Mg^{2+}$ of -5.7 and -5.9 kcal/mol. It is evident that interaction with $Mg^{2+}$ has increased the binding affinity towards substrates. Molecular dynamics of these complexes was carried out for 300ns using GROMACS 2022.2 software (Van Der Spoel et al., 2005). Further, each complex bound to substrates after a 300ns simulation was analyzed to validate the stable binding. The results revealed that in the absence of $Mg^{2+}$, F6P was not bound to the active pocket (**Figure 4A**), and S7P was bound to the active pocket (**Figure 4C**). Nevertheless, in the presence of $Mg^{2+}$, S7P formed more interactions with the catalytic site of PFP (**Figure 4D**). Also, a stable binding of F6P was observed in the presence of $Mg^{2+}$ in the active pocket (**Figure 4B**). Hence, for further analysis, the substrate bound with PFP+$Mg^{2+}$ was considered. The active site residues involved in the binding of S7P with PFP+$Mg^{2+}$ include G12, G13, R79, D122, D123, T124, T150, D152, R197 and R327. The substrate S7P formed polar contacts with the following residues: G13, I14, G121, D123, T124, K149, R327, and PPi, in the binding pocket, thereby strengthening the interaction (**Figure 4D**). The F6P formed interaction with the residues G12, G13, D122, T150, D152, M195 and R327, along with PPi, in the binding pocket of PFP+$Mg^{2+}$. Also, F6P forms polar contact with the residues, A11, G12, G13, I14, K149 and I151 (**Figure 4B).** The interaction study results infer catalytic residues involved in the binding of F6P and S7P. To assess the stable binding of the substrates, analysis of the simulation trajectory was carried out.

After molecular dynamics (MD) run for 300 ns, trajectories were analyzed to understand the conformational stability using RMSD, root mean square fluctuation (RMSF), radius of gyration (Rg), number of hydrogen bonds formed, and solvent accessible surface area (SASA) analysis. Each simulation run was for 300 ns. The RMSD plotted with respect to backbone atoms depicted the average values of 0.32 nm, 0.19 nm, 0.24 nm, and 0.2 nm obtained for apo, PFP+$Mg^{2+}$, F6P, and S7P, respectively (**Figure 5A)**. The RMSD plotted with respect to ligand atoms showed the stable binding of the substrates F6P and S7P with an average value of 0.34 nm and 0.29 nm, respectively (**Figure 5B)**. The SASA evaluates the exposed residues of the protein to the solvent molecules. The average values of SASA obtained for apo, PFP+$Mg^{2+}$, F6P, and S7P were 199.1 $nm^2$, 199.5 $nm^2$, 202.1 $nm^2$, and 196.4 $nm^2$, respectively, implying the stable formation of the complexes (**Figure 5C)**. RMSF is used to measure the fluctuation of individual residues of the protein to form a stable complex with the ligands. The average RMSF values of 0.17 nm, 0.15 nm, 0.16, and 0.14 nm were attained for apo, PFP+$Mg^{2+}$, F6P, and S7P, respectively (**Figure 5D)**. The RMSF plot shows that the ligands formed a stable complex with the protein by reducing the fluctuation in the loop regions and active pocket

residues. Rg values help to understand the compactness of the protein. The average Rg values obtained for apo, PFP+$Mg^{2+}$, F6P, and S7P were 2.45 nm, 2.42 nm, 2.41 nm, and 2.37 nm, respectively. Thus, the binding of the substrate didn't impact the packing of the protein (**Figure 5E)**. The number of hydrogen bonds measures the intermolecular contact formed between the atoms of protein and ligand to make the binding stronger. The average number of hydrogen bonds formed by F6P and S7P with PFP was 2.9 and 4.5, respectively (**Figure 5F)**. The clustering of the trajectory frames with respect to the RMSD of the substrate, using a cut-off of 0.15 nm for the last 20,000 frames of the 300 ns simulation, showed a single cluster for F6P and S7P, with average RMSD of 0.199nm and 0.188 nm, respectively (**Supplementary Figure 3**). The simulation results infer the stability of complexes formed with PFP by substrates, F6P, and S7P. Also, inferring that PFP can catalyze two reactions involved in the glycolytic and pentose phosphate pathway (PPP) (see section **2.4**).

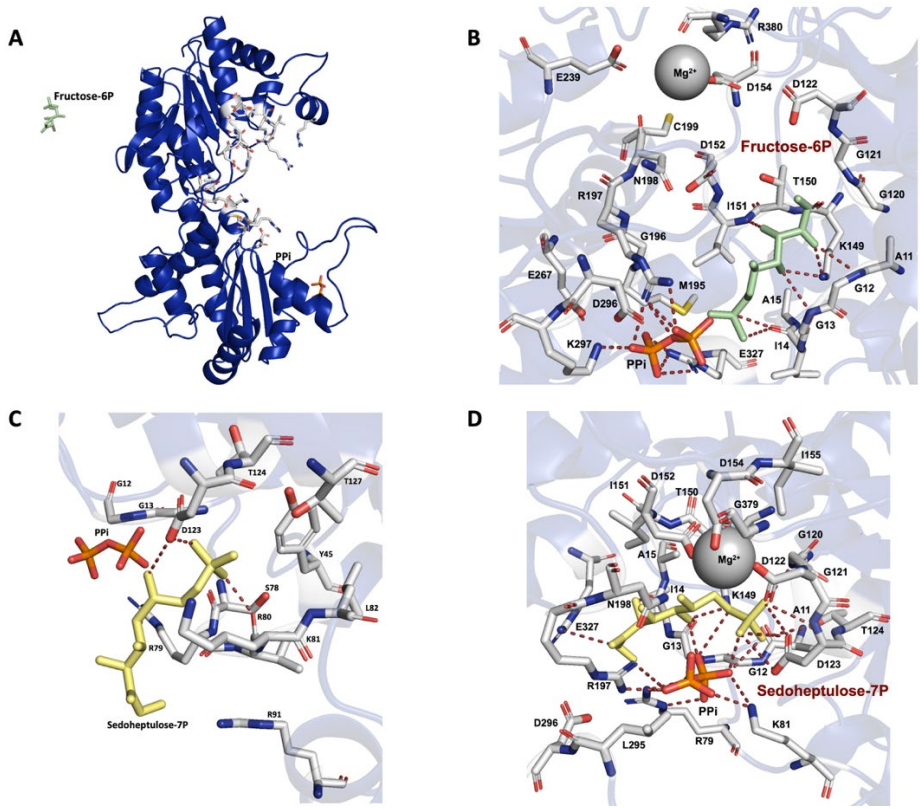


**Figure 4.** The PFP protein in complex with F6P (**A**) and S7P (**C**) in the absence of $Mg^{2+}$ in the active pocket after a simulation run of 300 ns. The PFP+$Mg^{2+}$ in complex with F6P (**B**) and S7P (**D**) showing an enlarged view of interactions formed after a simulation run of 300 ns. The hydrogen bond interactions are shown as red dotted lines.

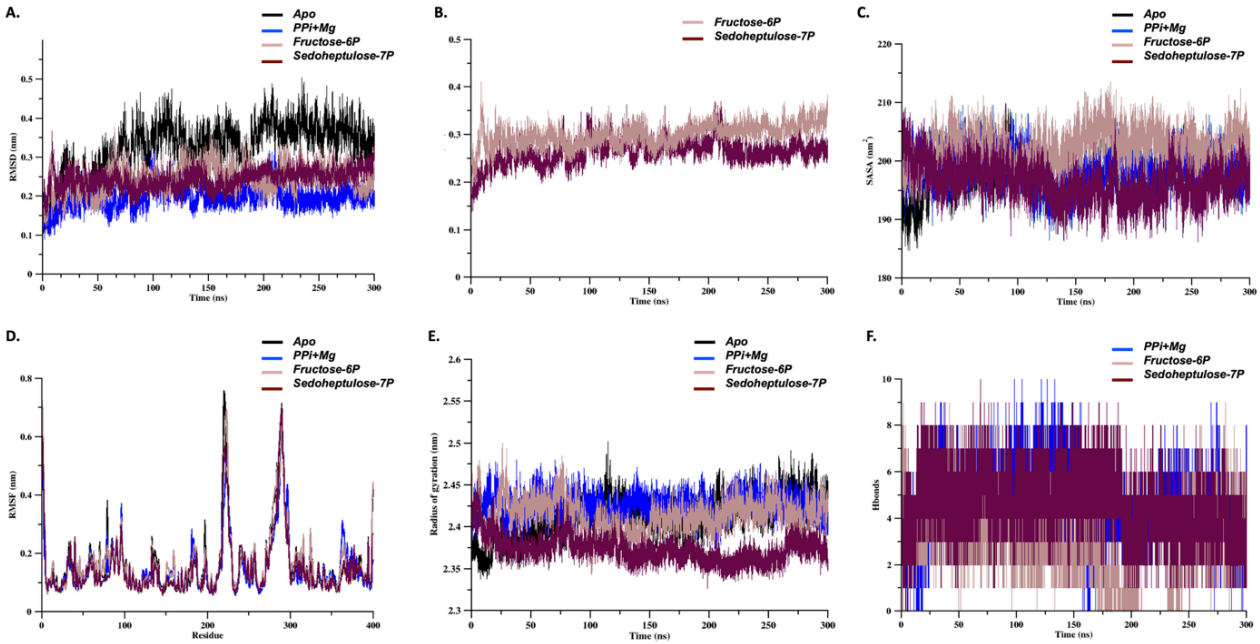


**Figure 5.** The structural stability analysis result after the simulation run for 300 ns. The plots of (**A**) RMSD-Backbone atoms, (**B**) RMSD-ligand atoms, (**C**) SASA, (**D**) RMSF, (**E**) Rg, and (**F**) number of H bond interactions represent the stability of PFP and complexes throughout the trajectories.

To obtain the effective free energy or enthalpy of binding ($\Delta G_{bind}$) of substrates with PFP+$Mg^{2+}$ model, MMGBSA and MMPBSA calculations were employed. The last 1000 frames from the MD run were utilized for the calculation by using the single trajectory protocol for multi-component system. In the study, the receptor considered was the PFP+$Mg^{2+,}$ and the ligand was the substrates (F6P and S7P). In which, $\Delta G_{bind}$ is the sum of molecular mechanical energy, $\Delta E_{MM,}$ and solvation energy ($\Delta G_{Solv}$). In the single trajectory protocol, internal bond energy turns zero, hence $\Delta E_{MM}$ will be the contribution of Van der Waal energy ($\Delta E_{VdW}$) and Electrostatic energy ($\Delta E_{Ele}$). $\Delta G_{Solv}$ is the summation of polar ($\Delta G_{GB/PB}$) and non-polar component ($\Delta G_{non\text{-}polar}$). $\Delta G_{bind}$ for the complexes PFP+$Mg^{2+}$- F6P and PFP+$Mg^{2+}$- S7P were -22.5 kcal/mol and -40.42 kcal/mol when computed using MMGBSA and -14.63 kcal/mol and -38.1 kcal/mol when using MMPBSA (**Table 2**). The results showed that S7P binds strongly to the receptor protein by forming more electrostatic interactions, and Van der Waal interactions of -55.91 kcal/mol and -30.68 kcal/mol, respectively, along with a polar solvation energy of 51.66 kcal/mol in MMGBSA studies. The MMPBSA calculation for S7P showed an electrostatic energy and solvation energy of -111.12 kcal/mol and 107.01 kcal/mol,

respectively. The MM(G/P)BSA studies showed the formation of stable complexes of PFP+$Mg^{2+}$- F6P and PFP+$Mg^{2+}$- S7P, with an energetically favorable enthalpy of binding, and in which S7P showed higher affinity with more stable interactions.

| | **Interaction energy: $\Delta E_{MM}$ (kcal/mol)** | | **Solvation energy: $\Delta G_{Solv}$ (kcal/mol)** | | **Effective free energy: $\Delta G_{bind}$ (kcal/mol)** |
|---|---|---|---|---|---|
| **Ligands** | Van der Waal energy, $\Delta E_{VdW}$ | Electrostatic energy, $\Delta E_{Ele}$/ $\Delta E_{Scf}$ | Polar component, $\Delta G_{GB/PB}$ | Non-polar component, $\Delta G_{non\text{-}polar}$ | Enthalpy, ΔH, (kcal/mol) |
| MMGBSA | | | | | |
| S7P | -30.68 | -55.91 | 51.66 | -5.48 | -40.42 |
| F6P | -25.08 | -30.49 | 37.6 | -4.53 | -22.5 |
| MMPBSA | | | | | |
| S7P | -30.76 | -111.12 | 107.01 | -3.22 | -38.1 |
| F6P | -25.16 | -61.18 | 74.52 | -2.81 | -14.63 |
| QM/MMGBSA | | | | | |
| S7P | -8.85 | -83.82 | 47.92 | -5.43 | -50.19 |
| F6P | -9.95 | -31.61 | 30.35 | -4.25 | -15.46 |

**Table 2.** The summary of effective free energy calculations for ligands, F6P and S7P, bound with PFP+$Mg^{2+}$ using MMGBSA, MMPBSA, and QM/MMGBSA.

The per-residue decomposition analysis predicted the residues involved in the stable interaction formed by the ligands with PFP+$Mg^{2+}$ (**Figure 6**). The substrate F6P formed interaction with the residues A11, G12, G13, I14, K149, T150, I151, M195 and R327. In which R327 showed the lowest energy contribution of -1.68 kcal/mol, followed by I151, I14, and K149 with average energy values of -1.42 kcal/mol, -1.3 kcal/mol, and -1.23 kcal/mol, respectively. Compared with F6P, S7P formed more number interactions by A11, G12, G13, I14, R79, G120, G121, T124, K149, T150, I51, R327, $Mg^{2+}$ and PPi. The average value of energy contributions shown by the residues K149> PPi> R79> I151> I14> T150> R327> G120 were -2.7 kcal/mol, -2.35 kcal/mol, -1.29 kcal/mol, -1.2 kcal/mol, -1.19 kcal/mol, -1.18 kcal/mol, -1.17 kcal/mol, and -1.03 kcal/mol, respectively. From the previous study, we know that the reaction mechanism observed for PFK involves the nucleophile attack of 1-OH group of F6P on the terminal phosphoryl group of ATP/PPi, thus facilitating substrate phosphorylation (Moore et al., 2002c). Further, on breaking the phosphoanhydride bond, terminal phosphate enters a pentacoordinate state stabilized by the residues R72, T125, G11, along with $Mg^{2+}$ (Shirakihara & Evans, 1988). The corresponding residues in CLas PFP include R79, T150, G13, and these align with the

result obtained from the decomposition analysis (**Figure 6**). The PFP protein of *Borrelia burgdorferi* (PDB ID:1KZH) proposed that K203 (corresponding to K149 in CLas PFP) plays an important role in the coordination of PPi and stabilization of the transition state in PFP proteins (Moore et al., 2002c). The interaction study showed that K149 forms a hydrogen bond with F6P and S7P (**Figure 4B, 4D**) and forms a stable interaction with energy of -1.23 kcal/mol and -2.7 kcal/mol, respectively (**Figure 6**). Also, in 1PFK, this residue is substituted with glycine at position 124, and this reaction mechanism involving lysine can be unique to PFP proteins. Other key residues involved in the binding of F6P in 1KZH were R431, which aligns with residue R327 of CLas PFP (Moore et al., 2002c). Also, the mutation study reveals the involvement of M169 in 1PFK, which aligns with M195 in CLas PFP (Shirakihara & Evans, 1988; X. Wang et al., 1998). The residues obtained from the per-residue decomposition analysis were utilized to define the QM region in QM/MMGBSA calculations. The molecule S7P showed the binding affinity of -50.19 kcal mol in QM calculation when compared to F6P with -15.46 kcal/mol (**Table 2**). The average value of electronic self-consistent field energy ($\Delta E_{Scf}$) for the QM region where catalysis occurs was -31.61 kcal/mol and -83.82 kcal/mol for F6P and S7P, respectively.

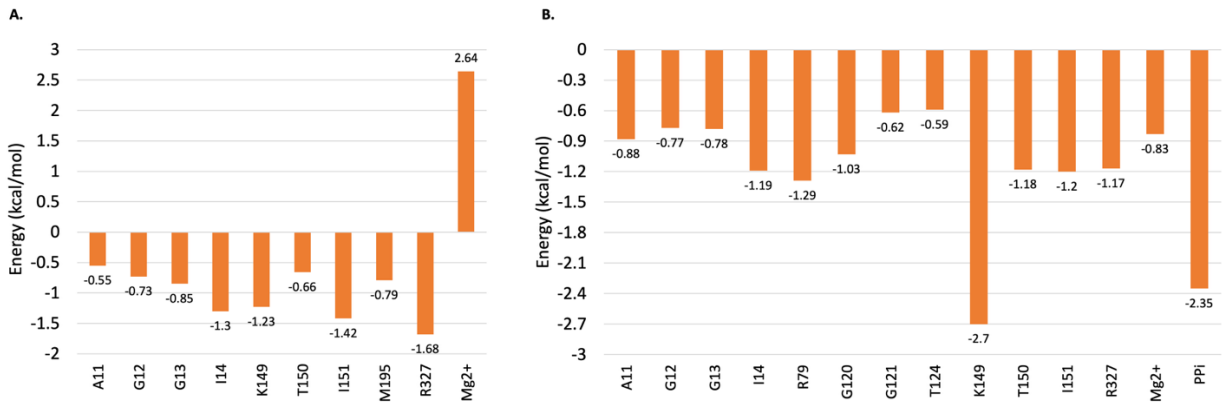


**Figure 6.** Per-residue decomposition analysis depicting the residues forming stable interaction with ligands A) F6P and B) S7P and corresponding energy contribution in kcal/mol.

### 2.4. Multifunctional role of PFP

Previous reports and sequence analysis revealed that CLas has an incomplete glycolytic pathway with the loss of *pgi* (**Figure 7: Step 2**) and PFK-1 (Attaran et al., 2020). In which PFK-1 catalyzes the reaction to convert F6P to fructose 1,6-bisphosphate (FBP) (**Figure 7: Step 13**). Studies support the expression of glycolysis-associated genes, except *pgi*, in planta, suggesting the utilization of glucose as the carbon source (Yan et al., 2013). One of the known

alternate route to complete the glycolytic pathway is using PPP (Stincone et al., 2015), which is disrupted in the absence of transaldolase enzyme (**Figure 7: Step 9**). The presence of functional PPP depending on the host (Zuñiga et al., 2020), indicates CLas might be using an alternate route to bypass and enter glycolysis. An alternate pathway has been reported in the human gut bacterium *Prevotella copri*, which utilizes the SBP pathway to metabolize C5 sugars and bypass the transaldolase step. The SBP pathway utilizes PFP protein to catalyze the reaction to produce SBP from S7P (**Figure 7: Step 10**) (Basen & Kurrer, 2021; Garschagen et al., 2021). This pathway variant is also found in *Entamoeba histolytica* and *cellulolytic clostridia* (Koendjbiharie et al., 2020; Susskind et al., 1982). In this study, we found through the structural analysis in section **2.3** that S7P has stable binding in the binding pocket of CLas PFP, revealing the presence of the active SBP pathway. Also, we show the binding of F6P with CLas PFP in section **2.3**. This implies that previously reported results about the disrupted glycolytic pathway in CLas can actually be circumvented by the dual role of CLas PFP in catalyzing both S7P (**Figure 7: Step 10**) and F6P (**Figure 7: Step 13**).

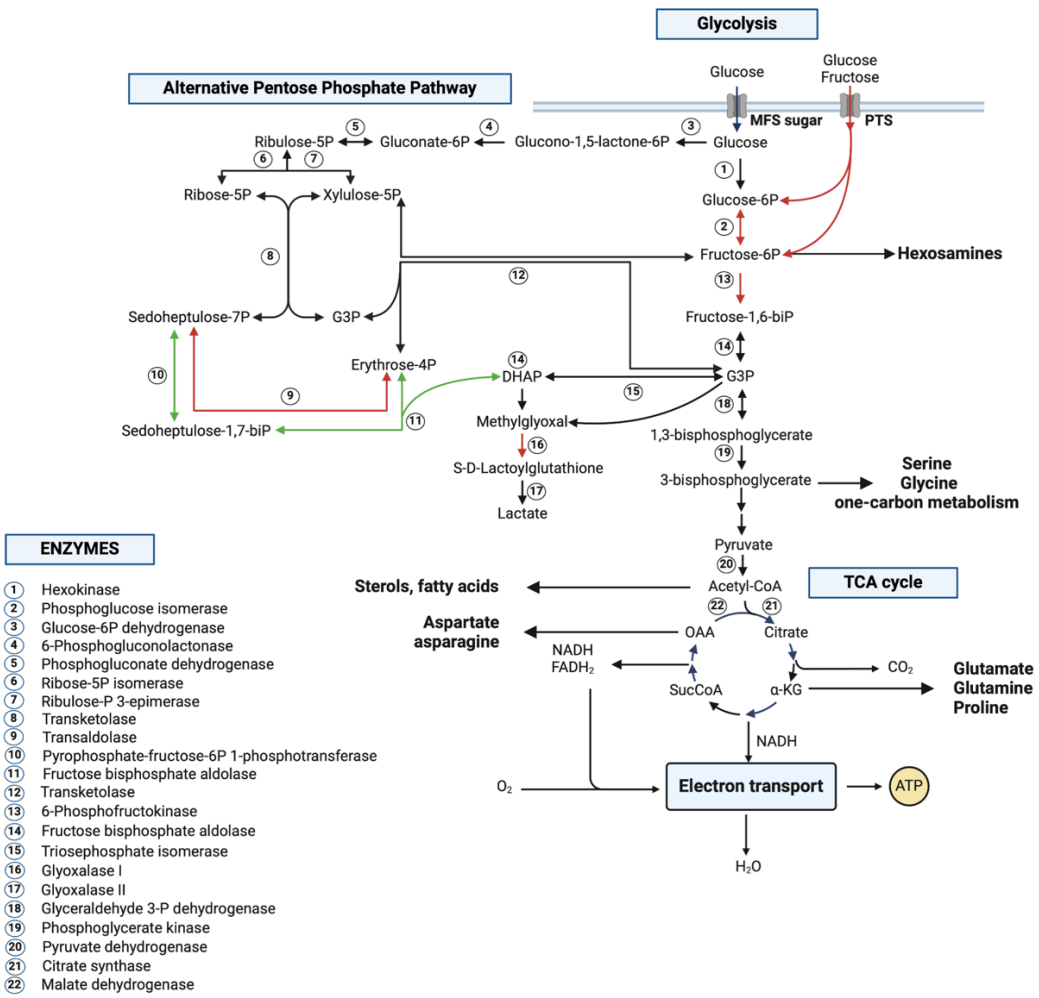


**Figure 7.** The representation of central carbon metabolism showing glycolysis, alternative pentose phosphate pathway, TCA cycle, Electron transport, and key enzymes. Genes present in CLas are indicated by black arrows, while absent genes are marked with red arrows. Alternative pathways are shown in green arrows. The metabolic pathway shows how PFP entering glycolytic pathway bypassing the loss of transaldolase (9) and can catalyze reaction for 6-phosphofructokinase (10), exhibiting a dual function. Created in BioRender. KP, S. (2025) https://BioRender.com/c31y167

Genome analysis studies have revealed the association between host-dependent obligate parasites and genome size reduction (Moran, 2002). CLas is known for its reduced genome and association with hosts (*citrus spp.* and *psyllid*) for metabolites for the survival by using transporters (Cai et al., 2022). The annotated transport proteins present in CLas to compensate for the missing genes are tabulated in **Supplementary Table 1**. Genome minimalism occurs in organisms for several reasons, including a) the ability to acquire metabolites directly from

the host, leading to the loss of biosynthetic genes/pathways; b) decrease in the necessity of the regulatory elements in a host-stable environment; c) genetic drift causing loss of functional genes; and d) the functional replacement by unrelated genes through non-orthologous gene displacement (Moran, 2002). The CLas PFP and FBPA (**Figure 7: Step 11**) were shown to replace the function of transaldolase. Hence, this raises the question whether this unusual intermediate reaction is an example of non-orthologous gene displacement (Koonin & Galperin, 2003). The reaction involving SBP as a metabolite is unusual, and the Calvin cycle in plants is the known pathway that involves SBP as an intermediate to form S7P using the enzyme sedoheptulose bisphosphatase (Ohta, 2022). FBPA in the Calvin cycle is a key enzyme involved in carbon fixation (T. Li et al., 2024), and the bifunctional role of FBPA showing activity towards FBP and SBP, has also been biochemically characterized (Flechner et al., 1999). Notably, a study involving transaldolase knockout mutants in *E.coli* showed the presence of an alternative pathway utilizing PFK-1 and accumulation of metabolites S7P in the absence of transaldolase along with SBP (Nakahigashi et al., 2009) (Fong et al., 2006). Hence, CLas might be using PFP to compensate for the lost metabolic reaction.

Another crucial enzyme absent in CLas is lactoylglutathione lyase gloA (EC 4.4.1.5), which is involved in the detoxification of the methylglyoxal formed during the glycolysis pathway (Jain et al., 2017) (**Figure 7: Step: 16**). It is unclear how CLas detoxifies methylglyoxal. However, the accumulation of methylglyoxal results from the following: unregulated uptake of sugar substrates other than glucose by sugar transporters and PTS system; increased synthesis of pentose sugars via transaldolase and transketolase; interruption of the glycolytic step producing 1,3-bisphosphoglycerate from G3P (**Figure 7: Step 18**) due to limited inorganic phosphate availability (Chakraborty et al., 2014). Hence, it is plausible that CLas may be regulating the methylglyoxal accumulation as it lacks gene *pgi*, PTS system, PFK-1, and transaldolase, reducing the uncontrolled production of sugar substrates and intermediates of glycolysis. Additionally, the reaction catalyzed by PFP in converting F6P to FBP and S7P to SBP aids in recycling inorganic phosphate from PPi, further mitigating metabolic imbalances by facilitating the conversion of G3P to 1,3-bisphosphoglycerate (**Figure 7: Step 18**)

### 2.5 Can CLas produce ATP?

Comparative sequence analysis of CLas against Lcr BT-1, the only culturable bacteria from *Liberibacter spp.,* was carried out to analyze the unique proteins/pathways present in CLas,

helping bacteria in energy production and survival. An incomplete glycolytic pathway can hinder ATP synthesis. So far in our analysis, we have unraveled the alternate pathway till the production of pyruvate, the product of glycolysis. The conversion of pyruvate to acetyl-CoA is a key metabolic step connecting glycolysis and the citric acid cycle or the TCA cycle. This step requires NAD+ and Coenzyme A (CoA) as co-substrates/cofactors. Pantothenate (Vitamin $B_5$) has a critical role in the biosynthesis of coenzyme A (CoA), but the reactions involved in the production of pantothenate are lost from CLas. The transporter for pantothenate is also not found in the annotated genome. However, CLas has the enzyme acyl-CoA thioesterase (EC 3.1.2.20) involved in detoxifying fatty acids to produce free fatty acids and CoA, which is absent in Lcr BT-1 (Park et al., 2021). It suggests a way of indirectly producing CoA and connecting the glycolysis with the TCA cycle (**Figure 7: Step 20**).

Lcr BT-1 and CLas, both species have cytochrome $bo_3$ (*cyoABCD*, EC 7.1.1.3) as a terminal oxidase (**Figure 8: cyt $bo_3$**). In the electron transport chain (ETC), electrons are transferred across ubiquinol and cytochrome $bo_3$ oxidase to oxygen, generating a proton gradient and leading to energy production via ATP synthase. But compared to Lcr BT-1, CLas lacks a complex of cytochrome bd oxidase (*cydX*) and cytochrome d ubiquinol oxidase (*cydB*) (**Figure 8: cyt bd**), which is used in oxygen detoxification and lacks a proton pump (Anand & Akhter, 2022; Richhardt et al., 2013). Cytochrome bd oxidase, a high-affinity terminal oxidase in *Escherichia coli*, plays a critical role in reducing oxidative stress (D'mello et al., 1996) and exhibits quinol peroxidase activity (Al-Attar et al., 2016). The absence of *cyt bd* in the CLas genome may render it more vulnerable to oxidative stress from oxygen and hydrogen peroxide ($H_2O_2$). Furthermore, studies suggest that *cyt $bo_3$*, a terminal oxidase, functions effectively at low oxygen levels (Lunak & Noel, 2015), aligning with the hypoxic (~1–3%) to anoxic conditions of plant phloem (Shahzad et al., 2023), implying that the ETC in CLas can remain functional by maintaining microaerophilic conditions. It is reported that alkaline pH in citrus phloem (7.3–8.5) and the psyllid midgut, and reduced oxygen tension during infection (Hijaz & Killiny, 2014; Molki et al., 2019). Thus, these conditions likely favor the axenic culture of CLas.

The evolution of alphaproteobacteria in response to early environmental factors has shaped their bioenergetic metabolic capabilities. In 2019, Degli *et al*. proposed that high-affinity oxidases evolved before low-affinity oxidases. Accordingly, high-affinity oxidase cyt bd and

intermediate-affinity oxidase *cyt bo3* in Lcr BT-1 suggest its early evolutionary divergence preceding CLas (Borisov et al., 2011; Degli Esposti et al., 2019).

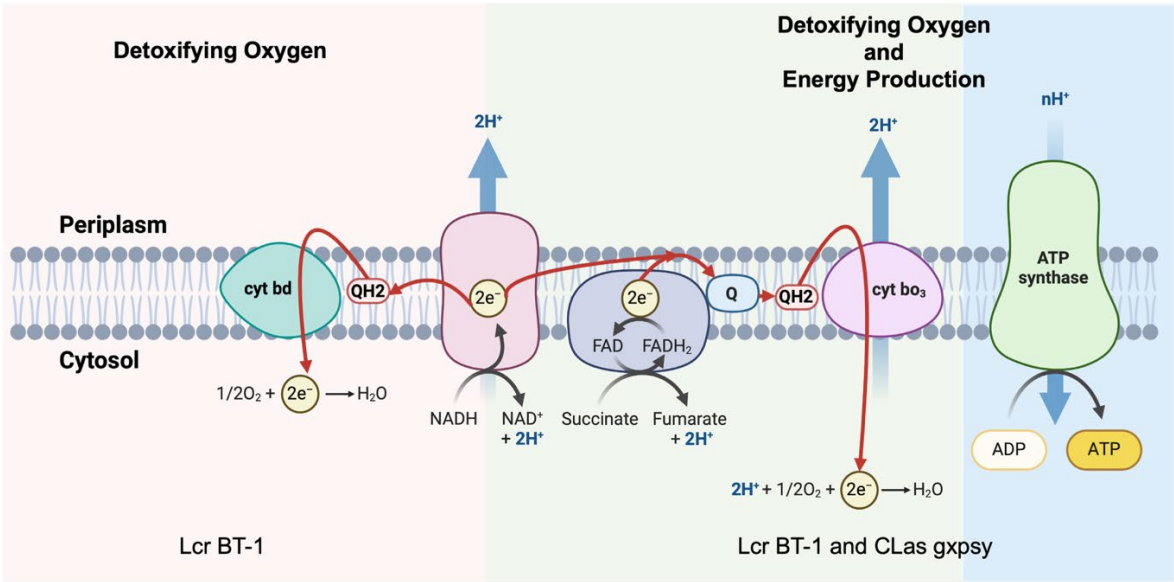


**Figure 8.** Schematic representation of the electron transport chain and energy production via *cyt bo₃* in CLas and Lcr BT-1, and the oxygen detoxification system through *cyt bd* in Lcr BT-1. Electron transfer across complexes is indicated by red arrows, while proton pumping is shown with blue arrows. Created in BioRender. KP, S. (2025) https://BioRender.com/g87z770

The presence of Npt1/Npt2 family nucleotide transporter enzyme in CLas, which is absent in Lcr BT-1, shows the dependence of CLas on the host for energy needs. Nevertheless, the reliance of CLas on the host for ATP and the downregulation of peroxidase activity in infected plants remain poorly understood (Pitino et al., 2017). In relation to that, *in vitro* studies have revealed that CLas infection enhances callose deposition on the phloem sieve pores of citrus plants, causing transport limitation and elevated $H_2O_2$ levels, intensifying HLB symptoms (Pitino et al., 2017; Welker et al., 2022). Thus, the host experiences a decline in ATP synthesis indirectly due to limited central carbon metabolism caused by restricted carbohydrate transport, forcing CLas to rely on its own mechanisms for ATP production. Additionally, CLas utilizes the secretory protein peroxiredoxin to reduce $H_2O_2$ level (Bernardini et al., 2022; Jain, Munoz-Bodnar, et al., 2019b). Moreover, studies showed that CLas suppresses the callose deposition in the seed vasculature of *Citrus sinensis* and *Citrus paradisi,* allowing its movement by increasing the sieve pore size (Bernardini et al., 2022; Welker et al., 2022).

Furthermore, CLas has a gene encoding 5/3'-nucleotidase survival protein (SurE), which is absent in Lcr BT-1. SurE is a stationary-phase survival protein found in eubacteria (except gram-positive bacteria) to maintain the homeostasis of nucleotides and phosphates, which is activated in stress-induced conditions, including nutrient deprivation, accumulation of reactive oxygen species, harsh pH and temperature, etc (Tarique et al., 2016; Zakataeva, 2021). Thus, we hypothesize that the CLas survives the limited nutrient availability caused by callose formation by enabling its transition to the stationary phase, activating the SurE gene, and suppressing plant defense responses. Also, CLas genome is encoded with Resistance to Homoserine/threonine family protein (RhtB) of LysE family translocators involved in the export of substrates including serine, homoserine, azaserine, leucine, threonine, alanine, cysteine, lactones, and 3-methylarginine across the membrane, creating a proton gradient (Tsu & Saier Jr, 2015). The possible reason for the export of these substrates could be to meet the requirement of the proton gradient and indirectly help with the ATP synthesis, or to regulate the metabolic flux and accumulation of amino acid-derived intermediates that enter the TCA cycle.

**Conclusion**

This study unravels the presence of PFP in CLas and the potential role in replacing the function of transaldolase. The catalytic domain involved in the binding of ATP/PPi in PFK sequences revealed distinct motifs with atypical amino acid variation. The phylogeny analysis showed early divergence of PFP compared with PFK-1 and PFK*, indicating PFP can be ancestral. Molecular clock hypothesis test failed for PFK* from *C.thermosuccinogenes,* inferring the less conserved ATP/PPi binding motif. The *in silico* structural studies employing molecular modeling and free energy calculations revealed the activity of the PFP protein, catalyzing two enzymatic reactions using F6P and S7P to sustain glycolysis. The role of $Mg^{2+}$ in regulating the strong binding of substrates to PFP was also assessed. The presence of *cyt* $bo_3$ and ATP:ADP translocase in the CLas shows the possible ways of acquiring energy substrates and the absence of terminal oxidase *cyt bd*, inferring the loss of the oxygen detoxification system. Additionally, the presence of *SurE* indicates a survival strategy under nutrient-limited conditions, potentially induced by host immune responses such as callose deposition. CLas exhibits a delicate balance between genome minimalism and functional adaptability, utilizing transport proteins and alternative biosynthetic pathways to acquire essential metabolites. For instance, despite losing genes for pantothenate biosynthesis, CLas possibly retains CoA using

acyl thioesterase. Further *in vitro* studies employing PFP as the target to control CLas can be studied. The critical analysis of the crucial or unique pathways will help to understand the survival mechanism in CLas and establish an axenic culture.

## 3. Methodology

### 3.1 Sequence analysis

The protein sequence of pyrophosphate fructose 6-phosphate 1-phosphotransferase (PFP) (WP_015452326.1) was retrieved from NCBI, and the sequence was subjected to the prediction of active site residues using NCBI conserved domain search (Marchler-Bauer et al., 2007). The sequence similarity search was performed using BLASTP search against the protein data bank database and the reference sequence proteins database. The multiple sequence alignment (MSA) was carried out for homologs selected based on the literature review using the MAFFT program (Katoh et al., 2002), utilizing the FFT-NS-i iterative refinement method and BLOSUM62 as the substitution matrix. The genome and coding sequence (CDS) of CLas strain gxpsy (NC_020549.1) and Lcr strain BT-1 (NC_019907.1) were fetched from Genbank, NCBI (Benson et al., 2012). The genome of the CLas and Lcr BT-1 were compared based on sequence similarity with an e-value cutoff of 1e-05 using protein-protein BLAST 2.14.0 software (Altschul et al., 1990) and assessed the unique and lost genes from CLas. The functional annotation of unique genes in CLas was carried using BLASTP by doing sequence similarity search against a database created from CDS sequences of the homologs from the phyla Proteobacteria, Firmicutes, Bacteroidota, and Mycoplasmatota.

### 3.2 Molecular phylogenetics

The molecular phylogenetic tree was constructed using aligned sequences of PFP and homologs using the maximum likelihood method in MEGA 11 software (Tamura et al., 2011) with 1000 replicates. To evaluate the consistent evolutionary rate among homologs of PFK, Tajima's relative rate test was conducted, and the molecular clock hypothesis was tested (Tajima, 1993)

### 3.3 Metabolic pathway analysis

The genome annotation of Lcr BT-1 and CLas was carried out using the RAST annotation server (Aziz et al., 2008), and the resulting genome sequence data were deployed to build a

metabolic model using Kbase software (Arkin et al., 2018). The metabolic model of CLas and Lcr BT-1 were compared to see the unique reactions in CLas as well as the lost metabolic functions. These results were compared with the functional annotation results and the pathways were manually curated. The annotation of transport genes in CLas was conducted by sequence and domain similarity search against sequence dataset obtain from Transport Classification Database (TCDB) (Saier et al., 2021).

### 3.4 Structure prediction

The three-dimensional structure of PFP protein found in CLas was predicted using AlphaFold3 (Abramson et al., 2024), RoseTTAFold server (Krishna et al., 2024), and ESMFold (Fang et al., 2023). In which PFP protein bound to $Mg^{2+}$ ion was modelled using AlphaFold3. The predicted structures were compared with ATP-dependent phosphofructokinase (PDB ID: 1PFK) from *E. coli* to select the best structure model. The modeled proteins were further validated by checking stereochemistry using PROCHECK (Laskowski et al., 1996) in SAVES server v6.0. The validated model was docked with cofactor pyrophosphate (PPi) using Autodock vina (Valdés-Tresanco et al., 2020) and was refined using molecular dynamics simulation in GROMACS 2022.2 software (Van Der Spoel et al., 2005) by CHARMM36 forcefield (Huang & MacKerell, 2013). The final model was visualized and analyzed in PyMOL software (Seeliger & de Groot, 2010). The three-dimensional structure of PFP protein from P.copri was also modeled using AlphaFold3 and refined using GROMACS software.

### 3.5 Molecular docking and structure refinement

The two-dimensional structures of substrates fructose 6-phosphate (F6P) and sedoheptulose 7-phosphate (S7P) were retrieved from PubChem (Q. Li et al., 2010). Both protein and ligand structures are pre-processed, and rigid docking was carried out using Autodock Vina in AMDock 1.5.2 graphical tool (Valdés-Tresanco et al., 2020). Based on the binding affinity, the best docked poses were selected, and interaction analysis was performed in PyMOL. Molecular dynamics of the protein-ligand complex (Lemkul, 2024) were carried out using GROMACS software. The best ligand poses obtained from docking were assigned with H atoms using Avogadro software (Hanwell et al., 2012). The ligand topology and parameterization were performed using the CHARMM General Force Field (CGenFF) webserver (Vanommeslaeghe et al., 2012). In contrast, the CHARMM forcefield and TIP3P water module of GROMACS were utilized to generate protein topology. The initial complex of the system was generated by combining protein (PFP), cofactor (PPi and $Mg^{2+}$), and respective substrate (F6P/S7P)

structures and corresponding topology were appended. A unit cell of a dodecahedron shape with a 2.0 nm padding centering protein-ligand complex was generated, and an SPC module was used to solvate the system. Further, the neutralization of the system was done by adding $Na^+/Cl^-$ ions and utilizing the steepest descent method to perform energy minimization for 100ps, converging to a maximum force of 1000 kJ/mol/nm. The position restraint topology was generated for the ligand structures and was applied while equilibrating the system. The NVT and NPT equilibration steps were carried out for 300 K and 1 bar pressure, respectively. Finally, the system was subjected to a 300 ns simulation run. The parameters used to elucidate structural stability include Root Mean Square Deviation (RMSD) with respect to backbone atoms and ligand atoms, Root Mean Square Fluctuation (RMSF), Hydrogen bond formation, Solvent Accessible Surface Area (SASA), and Radius of gyration (Rg). The clustering of the trajectory frames with respect to RMSD of substrate with cut-off of 0.1 nm, 0.15 nm, 0.2 nm for last 20,000 frames of 300 ns simulation was performed to obtain the consistent binding mode throughout the simulation.

### 3.6 Free energy calculation

To calculate the effective free energy of binding of PFP+$Mg^{2+}$-substrate complexes, Molecular Mechanic/Generalized Born Surface Area (MMGBSA), Molecular Mechanic/Poisson-Boltzmann Surface Area (MMPBSA) and Quantum Mechanics (QM/MMGBSA) method was used by gmx_MMPBSA v1.6.4 and AmberTools-20 (Valdés-Tresanco et al., 2021). The last 1000 frames of simulation run was taken to calculate the binding energy for a multi-component system with PFP+$Mg^{2+}$ as the receptor and substrate molecule as the ligand. In MMGBSA, igb=8 model was used and in MMPBSA, pbradii=7 (CHARMM radii) was considered. The per-residue decomposition analysis was performed along with MMGBSA method to predict the key residues involved in the binding. Further, obtained critical residues was taken to define qm region for QM/MMGBSA calculation using PM6 method.

**Acknowledgment**

SKP thanks Ministry of Education (MHRD), Government of India, for supporting with the fellowship. The authors are grateful to the computational facility provided by the Bioinformatics Center (BIC), supported by Department of Biotechnology (DBT), Government of India. The authors are thankful to the reviewers for the valuable suggestions and comments.

**Author Contributions**

SKP: Conceptualization, Methodology, Visualization, Data curation, Formal analysis, Writing – Original draft. JS & AKS: Conceptualization, Validation, Supervision, Writing – Review and Editing.

**References**


1. Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., Ronneberger, O., Willmore, L., Ballard, A. J., Bambrick, J., Bodenstein, S. W., Evans, D. A., Hung, C.-C., O'Neill, M., Reiman, D., Tunyasuvunakool, K., Wu, Z., Žemgulytė, A., Arvaniti, E., … Jumper, J. M. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, *630*(8016), 493–500. https://doi.org/10.1038/s41586-024-07487-w
2. Al-Attar, S., Yu, Y., Pinkse, M., Hoeser, J., Friedrich, T., Bald, D., & de Vries, S. (2016). Cytochrome bd Displays Significant Quinol Peroxidase Activity. *Scientific Reports*, *6*(1), 27631. https://doi.org/10.1038/srep27631
3. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, *215*(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2
4. Anand, P., & Akhter, Y. (2022). A review on enzyme complexes of electron transport chain from Mycobacterium tuberculosis as promising drug targets. *International Journal of Biological Macromolecules*, *212*, 474–494. https://doi.org/10.1016/j.ijbiomac.2022.05.124
5. Arkin, A. P., Cottingham, R. W., Henry, C. S., Harris, N. L., Stevens, R. L., Maslov, S., Dehal, P., Ware, D., Perez, F., Canon, S., Sneddon, M. W., Henderson, M. L., Riehl, W. J., Murphy-Olson, D., Chan, S. Y., Kamimura, R. T., Kumari, S., Drake, M. M., Brettin, T. S., … Yu, D. (2018). KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nature Biotechnology*, *36*(7), 566–569. https://doi.org/10.1038/nbt.4163
6. Attaran, E., Berim, A., Killiny, N., Beyenal, H., Gang, D. R., & Omsland, A. (2020). Controlled replication of ' *Candidatus Liberibacter asiaticus* ' DNA in citrus leaf discs. *Microbial Biotechnology*, *13*(3), 747–759. https://doi.org/10.1111/1751-7915.13531
7. Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., Formsma, K., Gerdes, S., Glass, E. M., Kubal, M., Meyer, F., Olsen, G. J., Olson, R., Osterman, A. L., Overbeek, R. A., McNeil, L. K., Paarmann, D., Paczian, T., Parrello, B., … Zagnitko, O. (2008). The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics*, *9*(1), 75. https://doi.org/10.1186/1471-2164-9-75

8. Bapteste, E., Moreira, D., & Philippe, H. (2003). Rampant horizontal gene transfer and phospho-donor change in the evolution of the phosphofructokinase. *Gene*, *318*, 185–191. https://doi.org/10.1016/S0378-1119(03)00797-2
9. Basen, M., & Kurrer, S. E. (2021). A close look at pentose metabolism of gut bacteria. *The FEBS Journal*, *288*(6), 1804–1808. https://doi.org/10.1111/febs.15575
10. Batarseh, T. N., Batarseh, S. N., Morales-Cruz, A., & Gaut, B. S. (2023). Comparative genomics of the Liberibacter genus reveals widespread diversity in genomic content and positive selection history. *Frontiers in Microbiology*, *14*. https://doi.org/10.3389/fmicb.2023.1206094
11. Benson, D. A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D. J., Ostell, J., & Sayers, E. W. (2012). GenBank. *Nucleic Acids Research*, *41*(D1), D36–D42. https://doi.org/10.1093/nar/gks1195
12. Bernardini, C., Turner, D., Wang, C., Welker, S., Achor, D., Artiga, Y. A., Turgeon, R., & Levy, A. (2022). *Candidatus* Liberibacter asiaticus accumulation in the phloem inhibits callose and reactive oxygen species. *Plant Physiology*, *190*(2), 1090–1094. https://doi.org/10.1093/plphys/kiac346
13. Borisov, V. B., Gennis, R. B., Hemp, J., & Verkhovsky, M. I. (2011). The cytochrome bd respiratory oxygen reductases. *Biochimica et Biophysica Acta (BBA) - Bioenergetics*, *1807*(11), 1398–1413. https://doi.org/10.1016/j.bbabio.2011.06.016
14. Bové, J. M. (2006). Huanglongbing: a destructive, newly-emerging, century-old disease of citrus. . *J Plant Pathol.* , *88*, 7–37. https://doi.org/https://apsjournals.apsnet.org/doi/pdf/10.1094/PHYTO-12-12-0331-RVW
15. Cai, L., Jain, M., Munoz-Bodnar, A., Huguet-Tapia, J. C., & Gabriel, D. W. (2022). A synthetic 'essentialome' for axenic culturing of 'Candidatus Liberibacter asiaticus.' *BMC Research Notes*, *15*(1), 125. https://doi.org/10.1186/s13104-022-05986-5
16. Chakraborty, S., Karmakar, K., & Chakravortty, D. (2014). Cells producing their own nemesis: Understanding methylglyoxal metabolism. *IUBMB Life*, *66*(10), 667–678. https://doi.org/10.1002/iub.1324
17. Chi, A., & Kemp, R. G. (2000). The Primordial High Energy Compound: ATP or Inorganic Pyrophosphate? *Journal of Biological Chemistry*, *275*(46), 35677–35679. https://doi.org/10.1074/jbc.C000581200
18. Compton, J. A., & Patrick, W. M. (2025). The more we learn, the more diverse it gets: structures, functions and evolution in the Phosphofructokinase Superfamily. *The Biochemical Journal*, *482*(9), 467–483. https://doi.org/10.1042/BCJ20253024

19. Davis, M. J., Mondal, S. N., Chen, H., Rogers, M. E., & Brlansky, R. H. (2008). Co-cultivation of ‘ *Candidatus* Liberibacter asiaticus’ with Actinobacteria from Citrus with Huanglongbing. *Plant Disease*, *92*(11), 1547–1550. https://doi.org/10.1094/PDIS-92-11-1547
20. Degli Esposti, M., Mentel, M., Martin, W., & Sousa, F. L. (2019). Oxygen Reductases in Alphaproteobacterial Genomes: Physiological Evolution From Low to High Oxygen Environments. *Frontiers in Microbiology*, *10*. https://doi.org/10.3389/fmicb.2019.00499
21. D’mello, R., Hill, S., & Poole, R. K. (1996). The cytochrome bd quinol oxidase in Escherichia coli has an extremely high oxygen affinity and two oxygen-binding haems: implications for regulation of activity in vivo by oxygen inhibition. *Microbiology*, *142*(4), 755–763. https://doi.org/10.1099/00221287-142-4-755
22. Fang, Y., Jiang, Y., Wei, L., Ma, Q., Ren, Z., Yuan, Q., & Wei, D.-Q. (2023). DeepProSite: structure-aware protein binding site prediction using ESMFold and pretrained language model. *Bioinformatics*, *39*(12). https://doi.org/10.1093/bioinformatics/btad718
23. Flechner, A., Gross, W., Martin, W. F., & Schnarrenberger, C. (1999). Chloroplast class I and class II aldolases are bifunctional for fructose-1,6-biphosphate and sedoheptulose-1,7-biphosphate cleavage in the Calvin cycle. *FEBS Letters*, *447*(2–3), 200–202. https://doi.org/10.1016/S0014-5793(99)00285-9
24. Fong, S. S., Nanchen, A., Palsson, B. O., & Sauer, U. (2006). Latent Pathway Activation and Increased Pathway Capacity Enable Escherichia coli Adaptation to Loss of Key Metabolic Enzymes. *Journal of Biological Chemistry*, *281*(12), 8024–8033. https://doi.org/10.1074/jbc.M510016200
25. Garschagen, L. S., Franke, T., & Deppenmeier, U. (2021). An alternative pentose phosphate pathway in human gut bacteria for the degradation of C5 sugars in dietary fibers. *The FEBS Journal*, *288*(6), 1839–1858. https://doi.org/10.1111/febs.15511
26. Ha, P. T., He, R., Killiny, N., Brown, J. K., Omsland, A., Gang, D. R., & Beyenal, H. (2019). Host-free biofilm culture of “Candidatus Liberibacter asiaticus,” the bacterium associated with Huanglongbing. *Biofilm*, *1*, 100005. https://doi.org/10.1016/j.bioflm.2019.100005
27. Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., & Hutchison, G. R. (2012). Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics*, *4*(1), 17. https://doi.org/10.1186/1758-2946-4-17
28. Hijaz, F., & Killiny, N. (2014). Collection and Chemical Composition of Phloem Sap from Citrus sinensis L. Osbeck (Sweet Orange). *PLoS ONE*, *9*(7), e101830. https://doi.org/10.1371/journal.pone.0101830

29. Huang, J., & MacKerell, A. D. (2013). CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *Journal of Computational Chemistry*, *34*(25), 2135–2145. https://doi.org/10.1002/jcc.23354
30. Jain, M., Cai, L., Fleites, L. A., Munoz-Bodnar, A., Davis, M. J., & Gabriel, D. W. (2019). *Liberibacter crescens* Is a Cultured Surrogate for Functional Genomics of Uncultured Pathogenic ' *Candidatus* Liberibacter' spp. and Is Naturally Competent for Transformation. *Phytopathology®*, *109*(10), 1811–1819. https://doi.org/10.1094/PHYTO-04-19-0129-R
31. Jain, M., Munoz-Bodnar, A., & Gabriel, D. W. (2017). Concomitant loss of the glyoxalase system and glycolysis makes the uncultured pathogen ``Candidatus Liberibacter asiaticus'' an energy scavenger. *Appl. Environ. Microbiol.*, *83*(23).
32. Jain, M., Munoz-Bodnar, A., & Gabriel, D. W. (2019). "Candidatus Liberibacter asiaticus" peroxiredoxin (LasBCP) suppresses oxylipin-mediated defense signaling in citrus. *J. Plant Physiol.*, *236*, 61–65.
33. John, P., & Whatley, F. R. (1975). Paracoccus denitrificans and the evolutionary origin of the mitochondrion. *Nature*, *254*(5500), 495–498. https://doi.org/10.1038/254495a0
34. Katoh, K., Misawa, K., Kuma, K.-I., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.*, *30*(14), 3059–3066. https://doi.org/https://doi.org/10.1093/nar/gkf436
35. Khademian, M., & Imlay, J. A. (2021). How microbes evolved to tolerate oxygen. *Trends Microbiol.*, *29*(5), 428–440. https://doi.org/https://doi.org/10.1016/j.tim.2020.10.001
36. Koendjbiharie, J. G., Hon, S., Pabst, M., Hooftman, R., Stevenson, D. M., Cui, J., Amador-Noguez, D., Lynd, L. R., Olson, D. G., & van Kranenburg, R. (2020). The pentose phosphate pathway of cellulolytic clostridia relies on 6-phosphofructokinase instead of transaldolase. *J. Biol. Chem.*, *295*(7), 1867–1878. https://doi.org/https://doi.org/10.1074/jbc.RA119.011239
37. Koonin, E. V., & Galperin, M. Y. (2003). Evolution of Central Metabolic Pathways: The Playground of Non-orthologous Gene Displacement. In *Sequence — Evolution — Function* (pp. 295–355). Springer US. https://doi.org/10.1007/978-1-4757-3783-7_8
38. Kricke, J., Mayer, F., Kopperschläger, G., & Kriegel, T. (1999). Phosphofructokinase-1 from Saccharomyces cerevisiae: analysis of molecular structure and function by electron microscopy and self-catalysed affinity labelling. *International Journal of Biological Macromolecules*, *24*(1), 27–35. https://doi.org/10.1016/S0141-8130(98)00063-4
39. Krishna, R., Wang, J., Ahern, W., Sturmfels, P., Venkatesh, P., Kalvet, I., Lee, G. R., Morey-Burrows, F. S., Anishchenko, I., Humphreys, I. R., McHugh, R., Vafeados, D., Li, X.,

Sutherland, G. A., Hitchcock, A., Neil Hunter, C., Kang, A., Brackenbrough, E., Bera, A. K., … Baker, D. (2024). Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science*, *384*(6693). https://doi.org/10.1126/science.adl2528

40. Kumar, P., Dalal, V., Kokane, A., Singh, S., Lonare, S., Kaur, H., Ghosh, D. K., Kumar, P., & Sharma, A. K. (2020). Mutation studies and structure-based identification of potential inhibitor molecules against periplasmic amino acid binding protein of Candidatus Liberibacter asiaticus (CLasTcyA). *International Journal of Biological Macromolecules*, *147*, 1228–1238. https://doi.org/10.1016/j.ijbiomac.2019.09.250
41. Kumar, P., Kesari, P., Kokane, S., Ghosh, D. K., Kumar, P., & Sharma, A. K. (2019). Crystal structures of a putative periplasmic cystine-binding protein from Candidatus Liberibacter asiaticus: insights into an adapted mechanism of ligand binding. *FEBS J.*, *286*(17), 3450–3472. https://doi.org/https://doi.org/10.1111/febs.14921
42. Laskowski, R. A., Rullmannn, J. A., MacArthur, M. W., Kaptein, R., & Thornton, J. M. (1996). AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *Journal of Biomolecular NMR*, *8*(4), 477–486. https://doi.org/10.1007/BF00228148
43. Lemkul, J. A. (2024). Introductory Tutorials for Simulating Protein Dynamics with GROMACS. *The Journal of Physical Chemistry B*, *128*(39), 9418–9435. https://doi.org/10.1021/acs.jpcb.4c04901
44. Leonard, M. T., Fagen, J. R., Davis-Richardson, A. G., Davis, M. J., & Triplett, E. W. (2012). Complete genome sequence of Liberibacter crescens BT-1. *Stand. Genomic Sci.*, *7*(2), 271–283. https://doi.org/https://doi.org/10.4056/sigs.3326772
45. Li, Q., Cheng, T., Wang, Y., & Bryant, S. H. (2010). PubChem as a public resource for drug discovery. *Drug Discov. Today*, *15*(23–24), 1052–1057. https://doi.org/https://doi.org/10.1016/j.drudis.2010.10.003
46. Li, T., Lin, Z., Zhu, C., Yang, K., Sun, H., Li, H., Wang, J., & Gao, Z. (2024). Identification and characterization of FBA genes in moso bamboo reveals PeFBA8 related to photosynthetic carbon metabolism. *International Journal of Biological Macromolecules*, *275*, 132885. https://doi.org/10.1016/j.ijbiomac.2024.132885
47. Lin, H., Lou, B., Glynn, J. M., Doddapaneni, H., Civerolo, E. L., Chen, C., Duan, Y., Zhou, L., & Vahling, C. M. (2011). The Complete Genome Sequence of 'Candidatus Liberibacter solanacearum', the Bacterium Associated with Potato Zebra Chip Disease. *PLoS ONE*, *6*(4), e19135. https://doi.org/10.1371/journal.pone.0019135

48. López, C., Chevalier, N., Hannaert, V., Rigden, D. J., Michels, P. A. M., & Ramirez, J. L. (2002). *Leishmania donovani* phosphofructokinase. *European Journal of Biochemistry*, *269*(16), 3978–3989. https://doi.org/10.1046/j.1432-1033.2002.03086.x
49. Lunak, Z. R., & Noel, K. D. (2015). A quinol oxidase, encoded by cyoABCD, is utilized to adapt to lower O2 concentrations in Rhizobium etli CFN42. *Microbiology*, *161*(Pt 1), 203–212. https://doi.org/https://doi.org/10.1099/mic.0.083386-0
50. Marchler-Bauer, A., Anderson, J. B., Derbyshire Myra K and DeWeese-Scott, C., Gonzales, N. R., Gwadz Marc and Hao, L., He, S., Hurwitz, D. I., Jackson, J. D., Ke, Z., Krylov, D., Lanczycki, C. J., Liebert, C. A., Liu, C., Lu Fu and Lu, S., Marchler, G. H., Mullokandov, M., Song, J. S., Thanki, N., Yamashita, R. A., … Bryant, S. H. (2007). CDD: a conserved domain database for interactive domain family analysis. *Nucleic Acids Res.*, *35*(Database issue), D237–40. https://doi.org/https://doi.org/10.1093/nar/gkl951
51. Martin, W. F., & Sousa, F. L. (2015). Early microbial evolution: The age of anaerobes. *Cold Spring Harb. Perspect. Biol.*, *8*(2), a018127. https://doi.org/https://doi.org/10.1101/cshperspect.a018127
52. Merfa, M. V., Pérez-López, E., Naranjo, E., Jain, M., Gabriel, D. W., & De La Fuente, L. (2019). Progress and Obstacles in Culturing ‘ *Candidatus* Liberibacter asiaticus’, the Bacterium Associated with Huanglongbing. *Phytopathology®*, *109*(7), 1092–1101. https://doi.org/10.1094/PHYTO-02-19-0051-RVW
53. Molki, B., Thi Ha, P., Mohamed, A., Killiny, N., Gang, D. R., Omsland, A., & Beyenal, H. (2019). Physiochemical changes mediated by ``Candidatus Liberibacter asiaticus'' in Asian citrus psyllids. *Sci. Rep.*, *9*(1), 16375. https://doi.org/https://doi.org/10.1038/s41598-019-52692-7
54. Montano, H. G., Bertaccini, A., & Fiore, N. (2024). Phytoplasma-Associated Diseases in South America: Thirty Years of Research. *Microorganisms*, *12*(7), 1311. https://doi.org/10.3390/microorganisms12071311
55. Moore, S. A., Ronimus, R. S., Roberson, R. S., & Morgan, H. W. (2002a). The Structure of a Pyrophosphate-Dependent Phosphofructokinase from the Lyme Disease Spirochete Borrelia burgdorferi. *Structure*, *10*(5), 659–671. https://doi.org/10.1016/S0969-2126(02)00760-8
56. Moore, S. A., Ronimus, R. S., Roberson, R. S., & Morgan, H. W. (2002b). The Structure of a Pyrophosphate-Dependent Phosphofructokinase from the Lyme Disease Spirochete Borrelia burgdorferi. *Structure*, *10*(5), 659–671. https://doi.org/10.1016/S0969-2126(02)00760-8

57. Moore, S. A., Ronimus, R. S., Roberson, R. S., & Morgan, H. W. (2002c). The structure of a pyrophosphate-dependent phosphofructokinase from the Lyme disease spirochete Borrelia burgdorferi. *Structure*, *10*(5), 659–671. https://doi.org/https://doi.org/10.1016/s0969-2126(02)00760-8
58. Moran, N. A. (2002). Microbial minimalism: genome reduction in bacterial pathogens. *Cell*, *108*(5), 583–586. https://doi.org/https://doi.org/10.1016/s0092-8674(02)00665-7
59. Murray, R. G., & Stackebrandt, E. (1995). Taxonomic note: implementation of the provisional status Candidatus for incompletely described procaryotes. *International Journal of Systematic Bacteriology*, *45*(1), 186–187. https://doi.org/10.1099/00207713-45-1-186
60. Nakahigashi, K., Toya, Y., Ishii, N., Soga, T., Hasegawa, M., Watanabe, H., Takai, Y., Honma, M., Mori, H., & Tomita, M. (2009). Systematic phenome analysis of *Escherichia coli* multiple-knockout mutants reveals hidden reactions in central carbon metabolism. *Molecular Systems Biology*, *5*(1). https://doi.org/10.1038/msb.2009.65
61. Noorizadeh, S., Golmohammadi, M., Bagheri, A., & Bertaccini, A. (2022). Citrus industry: Phytoplasma-associated diseases and related challenges for Asia, America and Africa. *Crop Protection*, *151*, 105822. https://doi.org/10.1016/j.cropro.2021.105822
62. Ohta, J. (2022). A novel variant of the Calvin–Benson cycle bypassing fructose bisphosphate. *Scientific Reports*, *12*(1), 3984. https://doi.org/10.1038/s41598-022-07836-7
63. Park, J., Kim, Y.-J., Lee, D., & Kim, K.-J. (2021). Structural basis for nucleotide-independent regulation of acyl-CoA thioesterase from Bacillus cereus ATCC 14579. *Int. J. Biol. Macromol.*, *170*, 390–396. https://doi.org/https://doi.org/10.1016/j.ijbiomac.2020.12.174
64. Pitino, M., Armstrong, C. M., & Duan, Y. (2017). Molecular mechanisms behind the accumulation of ATP and H2O2 in citrus plants in response to “Candidatus Liberibacter asiaticus” infection. *Horticulture Research*, *4*, 17040. https://doi.org/10.1038/hortres.2017.40
65. Rancati, G., Moffat, J., Typas, A., & Pavelka, N. (2018). Emerging and evolving concepts in gene essentiality. *Nat. Rev. Genet.*, *19*(1), 34–49. https://doi.org/https://doi.org/10.1038/nrg.2017.74
66. Richhardt, J., Luchterhand, B., Bringer, S., Büchs, J., & Bott, M. (2013). Evidence for a key role of cytochrome bo3 oxidase in respiratory energy metabolism of Gluconobacter oxydans. *J. Bacteriol.*, *195*(18), 4210–4220. https://doi.org/https://doi.org/10.1128/jb.00470-13

67. Saier, M. H., Reddy, V. S., Moreno-Hagelsieb, G., Hendargo, K. J., Zhang, Y., Iddamsetty, V., Lam, K. J. K., Tian, N., Russum Steven and Wang, J., & Medrano-Soto, A. (2021). The Transporter Classification Database (TCDB): 2021 update. *Nucleic Acids Res.*, *49*(D1), D461–D467. https://doi.org/https://doi.org/10.1093/nar/gkaa1004
68. Sechler, A., Schuenzel, E. L., Cooke, P., Donnua, S., Thaveechai, N., Postnikova, E., Stone, A. L., Schneider, W. L., Damsteegt, V. D., & Schaad, N. W. (2009). Cultivation of ' *Candidatus* Liberibacter asiaticus', ' *Ca.* L. africanus', and ' *Ca.* L. americanus' Associated with Huanglongbing. *Phytopathology®*, *99*(5), 480–486. https://doi.org/10.1094/PHYTO-99-5-0480
69. Seeliger, D., & de Groot, B. L. (2010). Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J. Comput. Aided Mol. Des.*, *24*(5), 417–422. https://doi.org/https://doi.org/10.1007/s10822-010-9352-6
70. Shahzad, F., Tang, L., & Vashisth, T. (2023). Unraveling the mystery of canopy dieback caused by citrus disease Huanglongbing and its link to hypoxia stress. *Frontiers in Plant Science*, *14*, 1119530. https://doi.org/10.3389/fpls.2023.1119530
71. Sharma, N., Selvakumar, P., Saini, G., Warghane, A., Ghosh, D. K., & Sharma, A. K. (2016). Crystal structure analysis in Zn 2+ -bound state and biophysical characterization of CLas-ZnuA2. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, *1864*(12), 1649–1657. https://doi.org/10.1016/j.bbapap.2016.08.016
72. Shirakihara, Y., & Evans, P. R. (1988). Crystal structure of the complex of phosphofructokinase from Escherichia coli with its reaction products. *Journal of Molecular Biology*, *204*(4), 973–994. https://doi.org/10.1016/0022-2836(88)90056-3
73. Siebers, B., Klenk, H.-P., & Hensel, R. (1998). PP i -Dependent Phosphofructokinase from *Thermoproteus tenax* , an Archaeal Descendant of an Ancient Line in Phosphofructokinase Evolution. *Journal of Bacteriology*, *180*(8), 2137–2143. https://doi.org/10.1128/JB.180.8.2137-2143.1998
74. Singh, A., Kumar, N., Tomar, P. P. S., Bhose, S., Ghosh, D. K., Roy, P., & Sharma, A. K. (2017). Characterization of a bacterioferritin comigratory protein family 1-Cys peroxiredoxin from Candidatus Liberibacter asiaticus. *Protoplasma*, *254*(4), 1675–1691. https://doi.org/10.1007/s00709-016-1062-z
75. Stincone, A., Prigione, A., Cramer, T., Wamelink, M. M. C., Campbell, K., Cheung, E., Olin-Sandoval, V., Grüning, N.-M., Krüger, A., Tauqeer Alam, M., Keller, M. A., Breitenbach, M., Brindle, K. M., Rabinowitz, J. D., & Ralser, M. (2015). The return of

metabolism: biochemistry and physiology of the pentose phosphate pathway. *Biol. Rev. Camb. Philos. Soc.*, *90*(3), 927–963. https://doi.org/https://doi.org/10.1111/brv.12140

76. Susskind, B. M., Warren, L. G., & Reeves, R. E. (1982). A pathway for the interconversion of hexose and pentose in the parasitic amoeba Entamoeba histolytica. *The Biochemical Journal*, *204*(1), 191–196. https://doi.org/10.1042/bj2040191
77. Tajima, F. (1993). Simple methods for testing the molecular evolutionary clock hypothesis. *Genetics*, *135*(2), 599–607. https://doi.org/10.1093/genetics/135.2.599
78. Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., & Kumar, S. (2011). MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, *28*(10), 2731–2739. https://doi.org/10.1093/molbev/msr121
79. Tarique, K. F., Abdul Rehman, S. A., Devi, S., & Tomar Priya and Gourinath, S. (2016). Structural and functional insights into the stationary-phase survival protein SurE, an important virulence factor of Brucella abortus. *Acta Crystallogr. F Struct. Biol. Commun.*, *72*(Pt 5), 386–396. https://doi.org/https://doi.org/10.1107/s2053230x16005999
80. Tsu, B. V, & Saier Jr, M. H. (2015). The LysE superfamily of transport proteins involved in cell physiology and pathogenesis. *PLoS One*, *10*(10), e0137184. https://doi.org/https://doi.org/10.1371/journal.pone.0137184
81. Vahling, C. M., Duan, Y., & Lin, H. (2010). Characterization of an ATP translocase identified in the destructive plant pathogen ``Candidatus Liberibacter asiaticus''. *J. Bacteriol.*, *192*(3), 834–840. https://doi.org/https://doi.org/10.1128/jb.01279-09
82. Valdés-Tresanco, M. S., Valdés-Tresanco, M. E., Valiente, P. A., & Moreno, E. (2020). AMDock: a versatile graphical tool for assisting molecular docking with Autodock Vina and Autodock4. *Biol. Direct*, *15*(1), 12. https://doi.org/https://doi.org/10.1186/s13062-020-00267-2
83. Valdés-Tresanco, M. S., Valdés-Tresanco, M. E., Valiente, P. A., & Moreno, E. (2021). gmx_MMPBSA: A New Tool to Perform End-State Free Energy Calculations with GROMACS. *Journal of Chemical Theory and Computation*, *17*(10), 6281–6291. https://doi.org/10.1021/acs.jctc.1c00645
84. Van Der Spoel, D., Lindahl, E., Hess, B., Groenhof, G., Mark, A. E., & Berendsen, H. J. C. (2005). GROMACS: fast, flexible, and free. *J. Comput. Chem.*, *26*(16), 1701–1718. https://doi.org/https://doi.org/10.1002/jcc.20291
85. Vanommeslaeghe, K., Raman, E. P., & MacKerell Jr, A. D. (2012). Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial

atomic charges. *J. Chem. Inf. Model.*, *52*(12), 3155–3168. https://doi.org/https://doi.org/10.1021/ci3003649

86. Wang, N., & Trivedi, P. (2013). Citrus huanglongbing: a newly relevant disease presents unprecedented challenges. *Phytopathology*, *103*(7), 652–665. https://doi.org/https://doi.org/10.1094/phyto-12-12-0331-rvw
87. Wang, X., Deng, Z., & Kemp, R. G. (1998). An Essential Methionine Residue Involved in Substrate Binding by Phosphofructokinases. *Biochemical and Biophysical Research Communications*, *250*(2), 466–468. https://doi.org/10.1006/bbrc.1998.9311
88. Wang, X., & Kemp, R. G. (1999). Identification of Residues of *Escherichia coli* Phosphofructokinase That Contribute to Nucleotide Binding and Specificity. *Biochemistry*, *38*(14), 4313–4318. https://doi.org/10.1021/bi982940q
89. Welker, S., Pierre, M., Santiago, J. P., Dutt, M., Vincent, C., & Levy, A. (2022). Phloem transport limitation in Huanglongbing-affected sweet orange is dependent on phloem-limited bacteria and callose. *Tree Physiol.*, *42*(2), 379–390. https://doi.org/https://doi.org/10.1093/treephys/tpab134
90. Yan, Q., Sreedharan, A., Wei, S., Wang, J., Pelz-Stelinski, K., Folimonova, S., & Wang, N. (2013). Global gene expression changes in Candidatus Liberibacter asiaticus during the transmission in distinct hosts between plant and insect. *Mol. Plant Pathol.*, *14*(4), 391–404. https://doi.org/https://doi.org/10.1111/mpp.12015
91. Zakataeva, N. P. (2021). Microbial 5’-nucleotidases: their characteristics, roles in cellular metabolism, and possible practical applications. *Appl. Microbiol. Biotechnol.*, *105*(20), 7661–7681. https://doi.org/https://doi.org/10.1007/s00253-021-11547-w
92. Zuñiga, C., Peacock, B., Liang, B., McCollum, G., Irigoyen, S. C., Tec-Campos, D., Marotz, C., Weng, N.-C., Zepeda, A., Vidalakis, G., Mandadi, K. K., & Borneman James and Zengler, K. (2020). Linking metabolic phenotypes to pathogenic traits among ``Candidatus Liberibacter asiaticus’’ and its hosts. *NPJ Syst. Biol. Appl.*, *6*(1), 24. https://doi.org/https://doi.org/10.1038/s41540-020-00142-w