Genomic Dissection of Metabolism, Prophage Elements, and Ecological Adaptations in *Candidatus* Liberibacter asiaticus

Sandra Kathott Prakash[1], Ashwani Kumar Sharma[1,*], Jitin Singla[1,2,*]

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

[2]Department of Computer Science and Engineering, 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 parasite that causes citrus greening disease in *citrus spp*. The effective measures to control the bacteria are hindered due to the lack of an established culturing method for the characterization of CLas. *Liberibacter crescens* (Lcr) is the only culturable model organism to study possible strategies to culture *Candidatus* Liberibacter species. The *in-silico* characterization of CLas was conducted through functional and comparative annotation of sequence data using sequence similarity search, metabolic pathway analysis, and prophage detection tools. The comparative genome analysis of CLas strain gxpsy and Lcr strain BT-1 showed a significant genome size reduction in the CLas genome and compensating for the lost essential genes with the help of transport proteins. The study focuses on the missing genes and the unique genes in CLas,

decoding the genomic and metabolic complexities. The study revealed the alternative pentose phosphate pathway utilized by the bacteria with the help of pyrophosphate fructose 6-phosphate 1-phosphotransferase (PPi-PFK) to bypass the loss of the transaldolase gene and to complete the glycolytic pathway. Also, the prophage genes are detected in the genome of CLas strains, possibly contributing to the genetic drift. This is the first report on the *in-silico* characterization of CLas exploring the survival strategies adopted by CLas in the metabolism to adapt to the host ecological conditions.

Keywords: citrus greening; glycolytic pathway; pyrophosphate fructose 6-phosphate 1-phosphotransferase; transaldolase; prophage.

1. Introduction

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 [1]. 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 [2,3]. The reactive oxygen species were initially detoxified using scavenging enzymes, including superoxide dismutases, peroxidases, and catalases [4]. Terminal oxidases were used by microbes for detoxification and later utilized oxygen for energy production [5].

Candidatus Liberibacter asiaticus (CLas) is a gram-negative, phloem-limited alphaproteobacteria transmitted by insect vector *Diaphorina citri*, commonly known as *psyllid*, causing the disease called “citrus greening” or Huanglongbing (HLB) in *citrus spp* [6] . Major symptoms shown by HLB-infected citrus plants are leaf spots, malformed/discolored fruits,

increased premature fruit abscission, and mortality rate [6]. HLB affects citrus production worldwide, causing a decline in the economy in the regions of Asia, Africa, Oceania, and the Americas [7]. An effective strategy to control the bacteria is unavailable due to lack of biophysical and functional characterization of CLas and its pathogenic mechanism.

Candidatus is a term used in prokaryotic taxonomy to describe bacteria or other prokaryotes that cannot be cultured. 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* equivalent to genus [8]. Various Liberibacter species that belong to the class alphaproteobacteria are included under *Candidatus* provisional status, including CLas, *Candidatus* Liberibacter africanus (CLaf), *Candidatus* Liberibacter americanus (CLam), *Candidatus* Liberibacter solanacearum (CLso), and *Liberibacter crescens* (Lcr).

Few bacteria under *Candidatus* status have been cultured over the years. For example, Lcr strain BT-1 was the first culturable bacteria of the genus Liberibacter spp., which was isolated from the infected leaf sap of Babaco Mountain papaya (*Carica stipulata* x *C. pubescens*) [9]. Lcr BT-1 is non-pathogenic and does not have any insect or plant host [10]. The axenic culture of Lcr BT-1 was grown using Basal Medium 7 (BM7) [11]. *Candidatus* Phytoplasma, a homologous to CLas, is a phloem-limited, mycoplasma-like organism that lacks a cell wall [12]. It has been reported to be associated with mixed infection with CLas and *Spiroplasma citri* to cause HLB in different citrus species [13]. It was cultured using a plant-based phytoplasma *in vitro* liquid medium [14,15].

Several trials for the pure culture of CLas have been made, but either it failed or can't maintain the viability of the bacteria over time. In 2008, Sechler *et al.* cultivated the pure culture of CLas using Liber A media composed of C3G and PD2 media by adding the composition to the citrus vein extract from petiole and midvein by maintaining microaerophilic conditions [16]. An attempt to culture CLas under host-free conditions has demonstrated that biofilms formed at lower oxygen tensions (10%) yield higher growth rates than aerobic and anaerobic conditions [17]. Co-culture experiments with actinobacteria related to *Propionibacterium acnes* also supported growth under microaerophilic conditions [18].

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 taken by the organism can either alter or lose these genes over time [19]. In the case of CLas, the prolonged association with the host leads to genome minimization, hence losing many essential genes [20]. CLas has 137 transport genes, of which 40 are ABC-type transporters capable of importing essential metabolites and thereby hampering the metabolic requirement of the citrus host, eventually leading to the death of the plant [21].

Previous studies have shown that CLas lacks essential enzyme genes throughout the glycolysis pathway, rendering the bacteria incapable of its own ATP production [22]. In 2010, Vahling *et al,* authors concluded that the ATP requirement of the bacteria is fulfilled by importing the ATP from the host using translocase (GU011685.1) [23]. In this study, we have shown that the CLas gxpsy can produce its own ATP by finding alternative metabolic reactions.

Genome studies have revealed the presence of prophage genes in the CLas genome. The first reported prophages were homologous to SC1 and SC2 phages [24] and Type-3 prophage was

detected in CLas strain JXGC [25]. The SC1-type phages are encoded with lytic genes integrated into the bacterial genome, and on the release of phage particles, citrus phloem cells can undergo apoptosis caused by the lytic burst of the CLas bacteria. The SC2-type phages, on the contrary, are temperate phages that take part in lysogenic conversion [24]. Fujiwara *et al.,* in 2018, reported that the growth of CLas strain ishi-1 is associated with the microbiota present in the citrus phloem. The growth density of CLas decreased with the treatment of oxytetracycline. However, the ishi-1 strain lacks a prophage region in the genome, and culturable methods were not applicable to other CLas strains [26].

This study compares the genome of the gxpsy strain of CLas with Lcr BT-1 to identify missing and alternative pathways or reactions that the bacteria incorporate for survival in the host. We identified an alternative route to complete the pentose phosphate pathway, bypassing the loss of the transaldolase enzyme, and validate it through *in silico* structural characterization. We also detect the presence of prophage genes in the CLas genome and analyze the order of their integration. Furthermore, we discuss the conditions and media components that could facilitate the axenic culture of CLas.

2. Results and Discussion

2.1 Genomic and Functional annotation

Genome analysis studies have revealed the association between host-dependent obligate parasites and genome size reduction [27]. 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; b) the redundancy of regulatory elements in a host-stable environment; c) genetic drift causing degradation or loss of essential genes; and d) non-

orthologous gene displacement, where unrelated genes perform equivalent metabolic functions [27]. Comparing the genome size of CLas gxpsy of ~1.28 mb with Lcr BT-1 of ~1.5 mb shows a significant reduction in genome size. CLas gxpsy has only 1078 protein-encoding genes compared to 1269 protein-encoding genes in Lcr BT-1. CLas has undergone evolutionary changes in its association with the host, leading to: loss of protein-coding genes and the partial or complete deletion of metabolic pathways; employing alternate routes or unique genes to complete the lost functions; genetic drift caused by prophage genes.

2.1.1 Metabolic pathway and 'missing links'

CLas is missing many essential genes required for glycolysis, methylglyoxal biosynthesis, amino acid synthesis, cofactor, prosthetic groups, carrier and vitamin biosynthesis, oxygen detoxification system, etc. (**Table 1**). The annotated transport genes present in CLas gxpsy to compensate for the missing genes are tabulated in **Table S1**. Here, we present a detailed analysis of missing genes and alternative pathways and reactions required for the survival of CLas, as well as other unique genes present in CLas.

Table 1. The number of genes present in CLas gxpsy and Lcr BT-1 against each metabolic pathway.

Metabolism		**No. of genes in CLas-gxpsy**	**No. of genes in Lcr-BT-1**
Carbohydrate metabolism	Glycolysis & gluconeogenesis	**12**	13
	Pentose phosphate pathway	8	**7**
	Pyruvate metabolism	5	5
	Citric acid cycle	15	15
	Methylglyoxal metabolism	**1**	3
Energy metabolism	Oxidative phosphorylation	**33**	36
Amino acid metabolism	Alanine, serine, and glycine	6	6
	Arginine; urea cycle, polyamines	9	9
	Aromatic amino acids and derivatives	**-**	19

	Glutamine, glutamate, aspartate, asparagine; ammonia assimilation	6	6
	Lysine, threonine, methionine, and cysteine	**16**	20
	Histidine	**-**	10
	Proline	**-**	4
Lipid metabolism	Fatty acid biosynthesis	13	13
	Fatty acid degradation	**-**	1
	Glycerolipid metabolism	**11**	12
	Acyl-CoA hydrolysis	**1**	-
Vitamin & cofactors, prosthetic groups, pigments	Thiamin	**2**	7
	Riboflavin	5	5
	Pyridoxal 5-phosphate	5	5
	Niacin	**5**	6
	CoA biosynthesis	**5**	7
	Biotin	5	5
	Folate	**8**	13
	Ubiquinone	5	5
Nucleotide metabolism	Purine	16	16
	Pyramidine	14	14

a) Glycolysis

On assessing central carbon metabolism and electron transport, we found that compared to Lcr BT-1, CLas gxpsy has an incomplete glycolytic (Embden-Meyerhof-Parnas) pathway and gluconeogenesis with the loss of glucose-6-phosphate isomerase (*pgi*) (EC 5.3.1.9) by which reversible interconversion of glucose-6-phosphate (G6P) and fructose-6-phosphate (F6P) is attained (**Figure 1: Step 2**). Adding to that, ATP-dependent phosphofructokinase (PFK-1) catalyzing the reaction to convert F6P to fructose 1,6-bisphosphate (FBP) is also absent in CLas and Lcr BT-1 (**Figure 1: Step 13**). The F6P is also not available directly through phosphoenolpyruvate-phosphotransferase system (PTS) as the CLas gxpsy lacks the PTS system. Although Lcr BT-1 can produce ATP using a direct glycolysis pathway, CLas gxpsy might be using alternate pathways to complete the glycolysis. Studies support the expression of glycolysis-associated genes, except *pgi*, in planta, suggesting the utilization of glucose as a carbon source [28].

The Pentose Phosphate Pathway (PPP) and Entner-Doudoroff (ED) pathway are two alternate pathways possible to complete glycolysis. ED requires two genes, namely, 6-phosphogluconate dehydratase, 2-keto-3-deoxyphosphogluconate aldolase, and both enzymes are missing in CLas gxpsy. The PPP is also not complete in CLas gxpsy. PPP consists of two phases. An irreversible oxidative phase that produces NADPH and D-ribulose 5-phosphate (Ru5P) from G6P (**Figure 1: Step 3-5**). And a reversible non-oxidative phase which uses Ru5P for the formation of 5-carbon (ribose 5-phosphate (R5P) and xylulose 5-phosphate (Xu5P)), 7-carbon (S7P), 4-carbon (erythrose 4-phosphate (E4P)), 6-carbon (F6P), and 3-carbon (glyceraldehyde 3-phosphate (G3P)) molecules **(Figure 1: Step 6-9, Step 12)** [29]**.** Sequence analysis inferred the absence of transaldolase (EC 2.2.1.2) enzyme reaction from the non-oxidative phase in CLas gxpsy (**Figure 1: Step 9**), hampering the synthesis of E4P. The absence of E4P disrupts the formation of G3P and F6P from Xu5P. Thus, the dependence on PPP to acquire the metabolites to complete the glycolysis pathway is also interrupted.

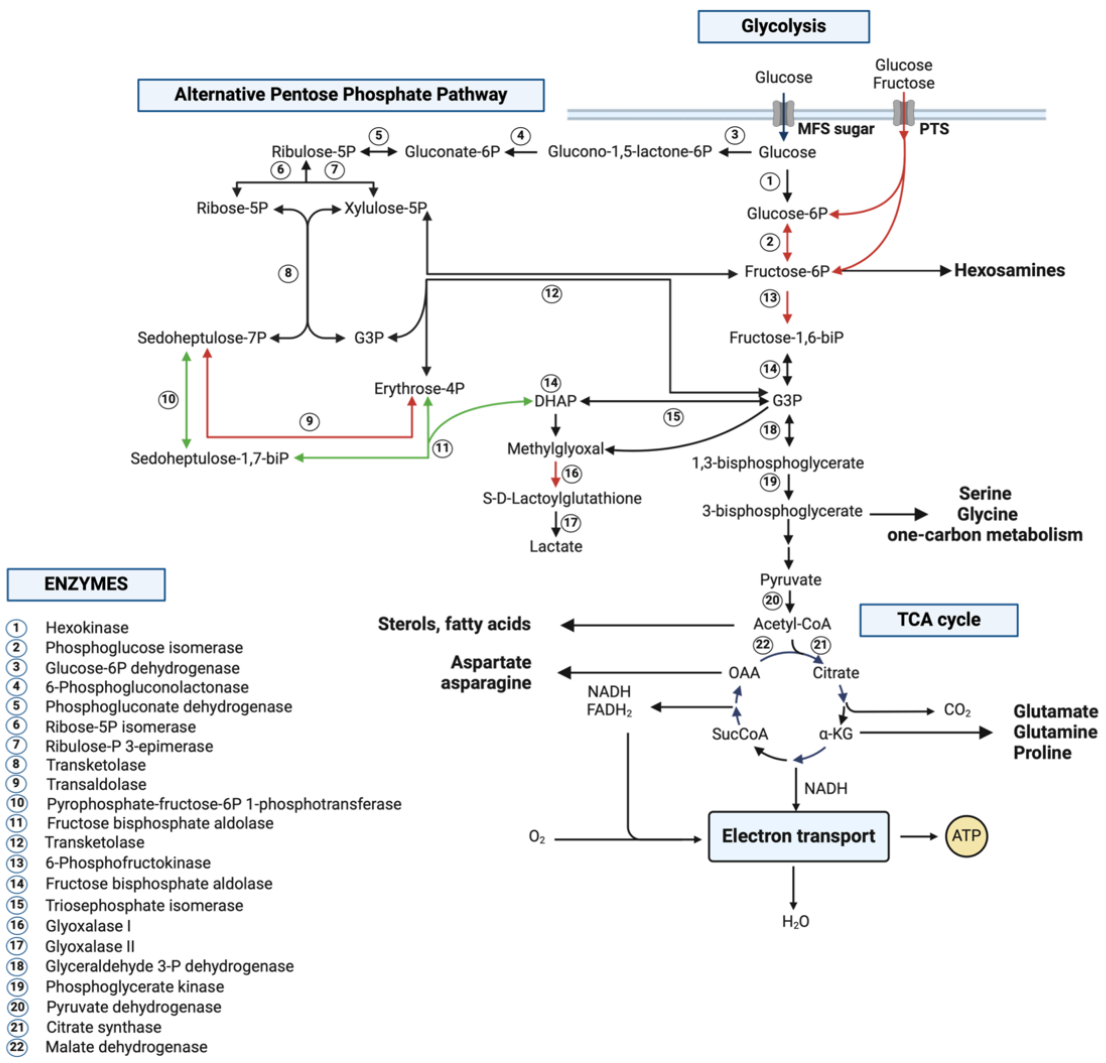

Figure 1. The representation of central carbon metabolism showing glycolysis, alternative pentose phosphate pathway, TCA cycle, Electron transport and key enzymes. Genes present in CLas gxpsy are indicated by black arrows, while absent genes are marked with red arrows. Alternative pathways are shown with green arrows. Created in BioRender. KP, S. (2025) https://BioRender.com/c31y167

Further analysis of metabolic pathways using KBase [30] shows the presence of a significant reaction catalyzed by pyrophosphate fructose 6-phosphate 1-phosphotransferase (PPi-PFK) (EC 2.7.1.90) to form D-sedoheptulose 1,7-bisphosphate (SBP) from S7P (**Figure 1: Step 10**). This reaction can evade the absence of transaldolase, thus completing PPP in an alternate route

(**Figure 1: Step 10-11**). A similar 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 [31,32]. This pathway variant is also found in *Entamoeba histolytica* and *cellulolytic clostridia* [33,34]. Flux distribution predictions using RNA sequencing data highlight reliance of CLas on the host for maintaining an active PPP, underscoring its ecological and energy dependence on the host environment [35]. In the later section **3.1.2**, we also show that the same enzyme can convert F6P to FBP (**Figure 1: Step 13**).

CLas gxpsy also lacks lactoylglutathione lyase gloA (EC 4.4.1.5), which is another crucial enzyme involved in the detoxification of the methylglyoxal formed during the glycolysis pathway [22] (**Figure 1: Step: 16**). However, it is unclear how CLas gxpsy is detoxifying methylglyoxal. The accumulation of methylglyoxal occurs when there is 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 glycolytic step producing 1,3-bisphosphoglycerate from G3P due to limited phosphate availability [36]. Thus, there is possibility that CLas may be evading the methylglyoxal accumulation as it lacks pgi, PTS system, PFK-1, and transaldolase, reducing the uncontrolled production of sugar substrates and intermediates of glycolysis. Additionally, the reaction catalyzed by PPi-PFK aids in recycling inorganic phosphate, further mitigating metabolic imbalances.

The oxidative carboxylation of pyruvate to form acetyl CoA and NADH is catalyzed by the enzyme pyruvate dehydrogenase complex (PDH) (EC 1.2.4.1) (Figure 1: Step 20). PDH catalyzes five consecutive reactions through substrate channeling by pyruvate dehydrogenase (E1), dihydrolipoamide transacetylase (E2), and dihydrolipoyl dehydrogenase (E3) [37][38]. The annotation of Pyruvate dehydrogenase, E1 component alpha subunit (pdhA) of CLas gxpsy

as a pseudogene due to the partial cleavage of the N-terminal of the protein showed a blockage to enter the tricarboxylic acid cycle (TCA). *In-silico* approach was deployed to study the structure and functional analysis of the protein pdhA. The active site residues involved in the binding of thiamine pyrophosphate (TPP) include Tyr51, Arg52, Gly100, Val102, Gly130, Asp131, Gly132, Gly133, Asn160, Tyr162, and His229. The structure of pdhA was predicted using AlphaFold v2.1.1 [39]**.** The sequence similarity search using BLASTP against PDB database predicted the homologous structures. The human pdhA protein (PDB ID: 6CFO) showed a higher percentage of identity and query coverage of 55.48% and 98%, respectively. The comparison of structures revealed a RMSD value of 0.58A°. The superimposed structures showed similar catalytic domains (**Figure 2**). Thus, implying the possibility of functional pdhA protein and CLas being able to enter the TCA cycle.

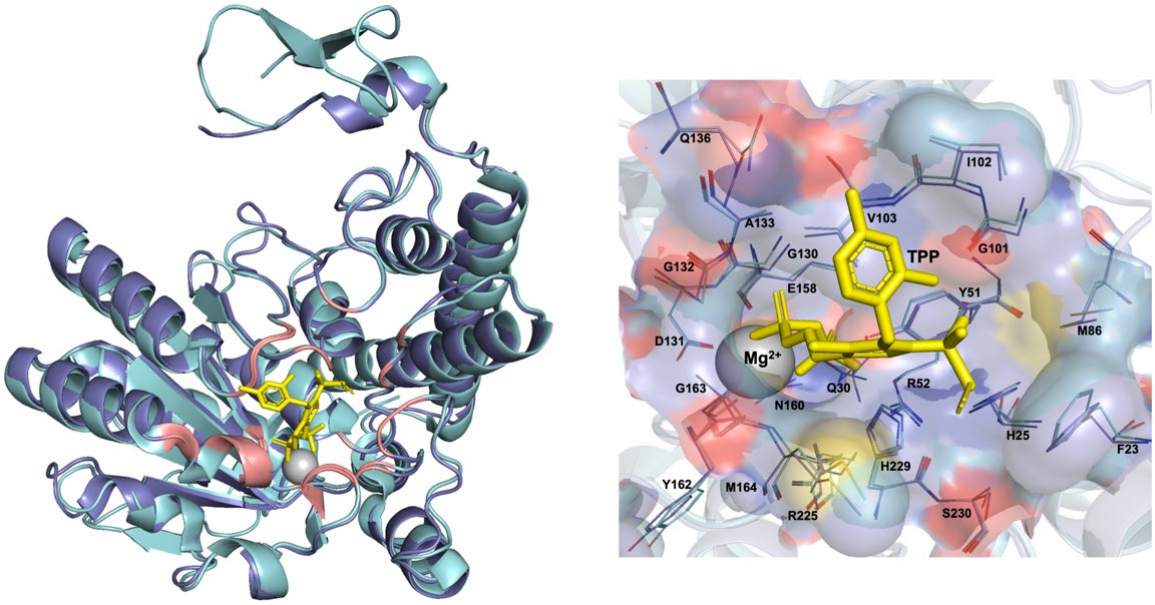

Figure 2. Superimposed structures of the predicted PdhA protein of CLas gxpsy (violet) and the PDB structure 6CFO (cyan) shown in cartoon representation. The catalytic domain bound with TPP (yellow stick) is shown in pink. The active site residues of pdhA protein of CLas gxpsy interacting with TPP is shown in the right side.

b) Energy production and survival

Lcr BT-1 and CLas gxpsy, both species have cytochrome bo_3, *cyoABCD* (EC 7.1.1.3) as a terminal oxidase (**Figure 3: 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 gxpsy lacks a complex of cytochrome bd oxidase (*cydX*) and cytochrome d ubiquinol oxidase (*cydB*) (EC 7.1.1.7) (**Figure 3: cyt bd**), which is used in oxygen detoxification and lack a proton pump [40]. Cytochrome bd oxidase, a high-affinity terminal oxidase in *Escherichia coli*, plays a critical role in reducing oxidative stress [41] and exhibits quinol peroxidase activity [42]. The absence of cydB and *cydX in* the CLas gxpsy genome may render it more vulnerable to oxidative stress from oxygen and hydrogen peroxide. Furthermore, studies suggest that *cyoABCD*, a terminal oxidase, functions effectively at low oxygen levels [43], aligning with the hypoxic (~1–3%) to anoxic conditions of plant phloem [44], implying that the ETC in CLas gxpsy remains functional. The citrus phloem's alkaline pH (7.3–8.5) and the psyllid vector's alkaline midgut pH, coupled with reduced oxygen tension during infection [45,46], likely create conditions that favor the axenic culture of CLas.

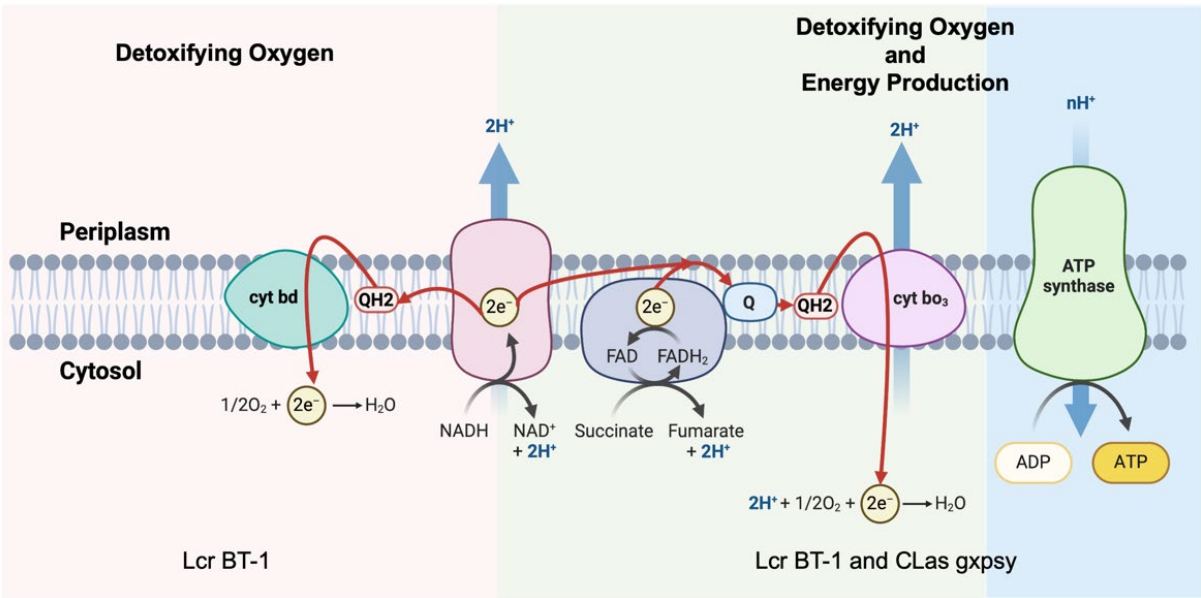

Figure 3. Schematic representation of the electron transport chain and energy production via cyt bo3 in CLas gxpsy and Lcr BT-1, and the oxygen detoxification system through cyt bd in Lcr BT-1. Electron transfer across complexes is indicated by red arrows, while proton pumping is shown with blue arrows. Created in BioRender. KP, S. (2025) https://BioRender.com/g87z770

The 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 cytochrome bd and intermediate-affinity oxidase *cyoABCD* in Lcr BT-1 suggest its early evolutionary divergence preceding CLas [5,47]. Comprehensive genome analysis further supports that Lcr BT-1 forms a basal lineage with Liberibacter spp., diverging as an outgroup [48].

Intriguingly, Npt1/Npt2 family nucleotide transporter enzyme was found in CLas gxpsy genome assigned for ATP:ADP antiporter activity, which is absent in Lcr BT-1. This indicates that CLas gxpsy relies on the host for the energy substrates. However, the dependence of CLas on the host for ATP and the downregulation of peroxidase activity in infected plants remain poorly understood [49]. CLas gxpsy genome is also 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 [50]. 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 access the intermediate substrates to enter the TCA cycle.

The 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 [49,51]. Thus, the host experiences a decline in ATP synthesis due to limited central carbon metabolism caused by restricted carbohydrate transport, forcing CLas to rely on its own mechanisms for ATP production. Moreover, it has been reported that CLas may suppress callose deposition in the seed vasculature of *Citrus sinensis* and *Citrus paradisi,* allowing its movement by increasing the sieve pore size. Additionally, CLas utilizes the secretory protein peroxiredoxin to reduce H_2O_2 level [52,53]. CLas gxpsy also has a gene encoding for 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 activates in stress-induced conditions including nutrient deprivation, accumulation of reactive oxygen species, harsh pH and temperature, etc. [54,55]. 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.

c) Amino Acid and Vitamins biosynthesis

CLas gxpsy and Lcr BT-1, both lack metabolic pathways to synthesize branched amino acids, including Leucine, Isoleucine, and Valine. This is compensated for by the presence of branched-chain amino acid ABC transporters. Lcr BT-1 and CLas gxpsy genome also lacks reactions to synthesize methionine consisting of homoserine O-acetyltransferase metA (EC 2.3.1.31), cystathionine γ-lyase metC (EC 4.4.1.1), methionine synthase metE or metHL (EC 2.1.1.13) except for the presence of cystathionine γ-synthase metB (EC 2.5.1.48) in Lcr BT-1. This is compensated with methionine ABC transporter ATP-binding protein (EC 7.4.2.11) in Lcr BT-1 and an ATP-dependent ABC transporter for the D-methionine in CLas gxpsy.

The genes associated with the amino acid metabolism, including biosynthesis of histidine, tryptophan, tyrosine, phenylalanine, cysteine, asparagine, and proline, were missing from CLas genome compared to Lcr BT-1. All these are also compensated by other transporters. CLas gxpsy has three different amino acid transporters: amino acid ABC transporter permease, branched-chain amino acid transporter, and dicarboxylate/amino acid: cation symporter. It has been reported that the alternate mechanism of putative cystine binding protein from CLas strain psy62 showed a broad specific affinity towards amino acids other than cationic amino acids [56].

The vitamins are essential for bacterial growth and survival and are involved in the metabolic pathways as cofactors. TPP is a crucial coenzyme that contributes to central carbon metabolism by taking part in the formation of acetyl-CoA from pyruvate and succinyl CoA from alpha-ketoglutarate [57]. CLas gxpsy has lost gene encoding for the biosynthesis of thiamine (Vitamin B_1), but the essentiality is maintained using thiamine ABC transporter permease ThiP ThiQ ThiB and thiamine diphosphokinase (EC 2.6.7.2) for transporting thiamine from the host and phosphorylation, respectively. 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 not found in the annotated genome. The CLas gxpsy has 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 [58]. It suggests a way of indirectly producing CoA and reducing fatty acid transformations. The transcriptomic and metabolic analysis of CLas-infected *D.citri* revealed a downregulation of genes associated with pantothenate biosynthesis and a significant reduction in fatty acid derivatives, aligning with our findings [59].

The shikimate pathway is entailed for the synthesis of chorismate, which is missing from CLas except for 3-phosphoshikimate 1-carboxyvinyltransferase (EC 2.5.1.19) associated with the formation of 5-enolpyruvoyl shikimate 3-phosphate from shikimate 3-phosphate. Chorismate is a significant metabolite that takes part in the biosynthesis of aromatic amino acids and folate (Vitamin B_9) by utilizing para-aminobenzoic acid (PABA) and dihydropteroic acid as intermediates. Amino Deoxy Chorismate Synthase (EC 2.6.1.85) engaged in producing PABA and dihydropteroate synthase (EC 2.5.1.15) for dihydropteroic acid are absent in CLas. The amino acid ABC transporter permease can be utilized to transport aromatic amino acids. Even though the genes involved in the synthesis of tetrahydrofolate and folate polyglutamylation are present, transporter or any other genes to maintain the essentiality of chorismate and PABA are not found. Liu *et al.,* in 2021, showed a significant reduction of folate formation using transcriptomic analysis, suggesting the potential modulation of vitamin and cofactor synthesis due to the association of CLas with the host [59].

The enzyme assigned as bifunctional nicotinamidase/pyrazinamidase from the NAD salvage pathway is also missing in CLas. Thus, the genes for the de novo synthesis of the nutrients by CLas were eliminated, and the bacteria solely relies on the host for these metabolites. The presence/absence of pathways associated with the synthesis of amino acids and vitamins in CLas and Lcr BT-1 are illustrated in **Figure 4**.

	CLas gxpsy	Lcr BT-1
Gly		
Leu	Branched-chain amino acid ABC transporter	Branched-chain amino acid ABC transporter
Ile	Branched-chain amino acid ABC transporter	Branched-chain amino acid ABC transporter
Val	Branched-chain amino acid ABC transporter	Branched-chain amino acid ABC transporter
Ala		
Met	ABC transporter ATP-binding	Methionine ABC transporter permease
Phe	Amino acid ABC transporter permease	
Trp	Amino acid ABC transporter permease	
Tyr	Amino acid ABC transporter permease	
Pro	ABC transporter permease subunit	
His	Amino acid ABC transporter permease	
Asp		
Glu		
Lys		
Arg		
Ser		
Thr		
Cys	Amino acid ABC transporter permease	
Asn	Amino acid ABC transporter permease	
Gln		
Biotin		
Folate	?	
Niacin		
Choline	ABC transporter ATP-binding	Choline ABC transporter ATP-binding
Thiamine	Thiamine ABC transporter permease	
Riboflavin		
Pyridoxine		
Pantothenate	?	
Ubiquinone		

Figure 4. The pathways present in CLas gxpsy and Lcr BT-1 for amino acid and vitamin biosynthesis are highlighted in green boxes, while missing pathways are indicated in red boxes, along with the identified transport proteins.

2.1.2 Prophages and genome rearrangement

Phages play a role in lysogenic conversions, potentially transducing genes that inactivate beneficial host bacterial genes and contributing to genome rearrangement and genetic drift [60,61]. The absence of essential genes in CLas gxpsy compared to Lcr BT-1 may be due to the presence of phage genes in CLas, which are absent in Lcr BT-1. Integrated prophages in the CLas genome influence its virulence. For example, peroxidase encoded in the prophage region can suppress HLB symptoms [62], whereas the expression of LasP235 from the

prophage region induces HLB-like symptoms in transgenic Carrizo citrange [63]. The study revealed the presence of lytic genes, including endolysin, holin, and peroxidase, integrated into the CLas genome.

SC1 and SC2 genomes share 93% identity among the early genes, and late genes are mostly unique [24]. The sequence similarity search revealed that SC1 has 7 unique genes (lytic genes) compared to SC2, and SC2 has 10 unique genes (lysogenic genes) compared to SC1 (**Figure 5A, 5B)**. We studied the presence of lytic and lysogenic genes in various *Candidatus* Liberibacter spp, using progressive alignment using Mauve software [64] and sequence similarity search using BLASTP. Significant findings include (a) the presence of SC2 lysogenic genes in the citrus-isolated CLas strains A4, GDCZ, GDCH12, TaiYZ2; (b) the presence of SC1 lytic genes in citrus-isolated CLas genome strains 9PA and YNBC and psyllid-isolated CLas strains gxpsy, CoFLP, psy62, and ReuSP1, except ishi-1, and JRPAMB1; (c) CLas gxpsy has both SC1 and SC2 types of prophages integrated into its genome sequence (**Figure 5C)**. The results indicate that lytic genes are primarily found in psyllid-isolated strains rather than citrus-isolated ones. Also, other orthologous species, including CLam, CLaf, and CLso, have SC1 and SC2 type prophage genes but lack lytic and lysogenic genes. To detect the prophage region in the bacterial genome of Liberibacter spp., DBSCAN-SWA tool [65] was utilized. The predicted prophage regions in the Liberibacter genome are shown in **Table S2.** The conserved domain search found a CRISPR associated exonuclease Cas4 (PD-(D/E) XK nuclease) protein (WSI_05335, WSI_05555, WSI_05755) in the prophage region of CLas gxpsy. It has been hypothesized that the integrated Cas protein can prevent the invasion of other phages to maintain single prophage-encoded bacterial strains [66]. Also, ishi-1 is the only CLas strain with a fragmented phage genome encoded for 7 prophage genes located at genome position 1184978-1190852 and has a successful axenic culture reported earlier [26]. Lcr BT-1's

culturability under aerobic conditions and independence from a host correlate with its lack of prophages [10]. Variations in phage components among host-specific CLas strains remain uncharacterized, highlighting the need for functional studies on their lytic and lysogenic genes.

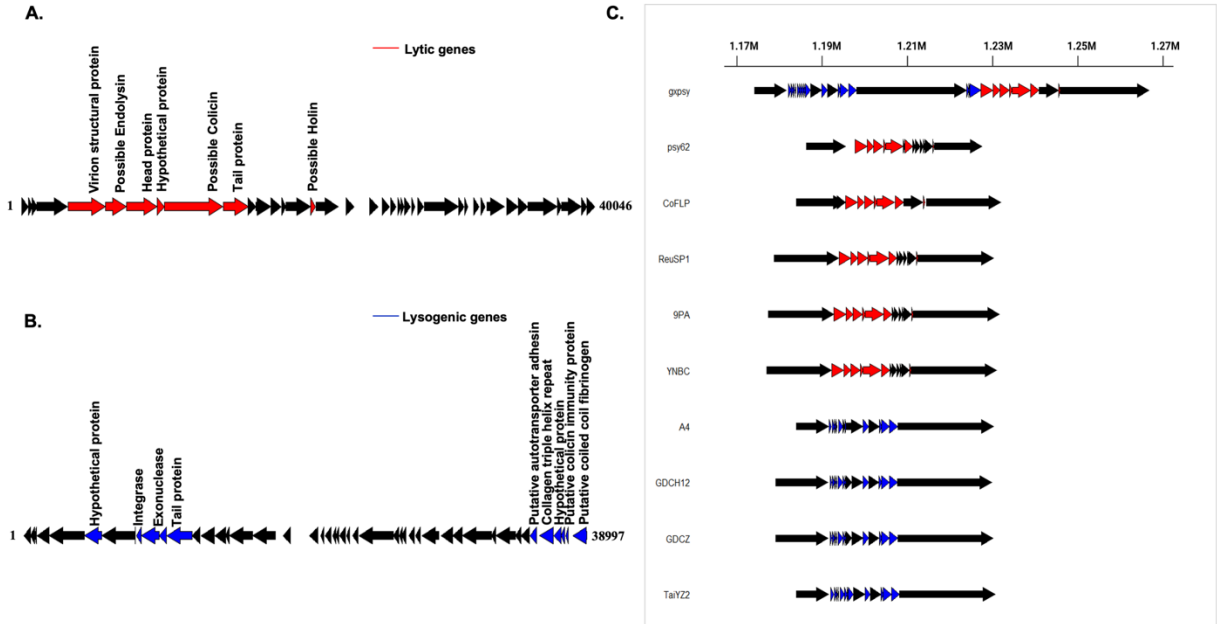

Figure 5. Schematic representation of the prophage genome. (**A**) The SC1 prophage genome, with lytic genes represented by red arrows. (**B**) The SC2 prophage genome, with lysogenic genes shown in blue arrows. Similar genes shared between SC1 and SC2 prophages are indicated by black arrows. (**C**) The presence of SC1 or SC2-type prophage genes in CLas strains is identified and plotted.

2.2 Molecular modeling of PPi-PFK

An unusual incorporation of intermediate reaction by PPi-PFK could be an alternative pathway that could bypass the lack of transaldolase. Microorganisms lacking transaldolase may rely on the SBP pathway due to the reaction's reversibility compared to PFK-1 and also helps in

conserving energy substrates in ATP-limited environment [33]. The sequence similarity search revealed that CLas only has a single enzyme, PPi-PFK, which plays an essential role in the glycolytic pathway and belongs to the phosphofructokinase type A (IPR050929) known as PFKA. The bioinformatics tool ProtParam inferred stable protein with theoretical pI, extinction coefficient, and molecular weight of 6.56, 34755M^{-1} cm^{-1}, and 44.4kDa, respectively. The protein structure was predicted using AlphaFold, ESMFold [67], and RoseTTAFold [68], and a comparison of predicted structures revealed the overlap in the catalytic domain shown in **Figure 6A**. The RMSD values calculated for the aligned structures on comparing with the catalytic site of PFK-1 from *E. coli* (PDB ID: 1PFK) were 0.63 A°, 2.71 A°, and 0.95 A° for AlphaFold, ESMFold, and RoseTTAFold, respectively. For further study, the structure predicted using Alphafold was used based on the structural similarity with 1PFK. The crystal structure of PFK-1 from *E. coli* showed the conformational change occurs on the phosphoryl transfer to the F6P and bound with Mg^{2+} in the active site [69]. The AlphaFold model comprises α/β domain forming fold to accommodate substrate and inorganic phosphate (PPi) and Mg^{2+}(PPi-PFK+Mg^{2+}). The active site residues predicted using the conserved domain search tool are Gly13, Tyr45, Arg79, Asp122, Asp123, Thr124, Thr126, Thr127, Thr150, Asp152, Asp154, Met195, Gly196, and Arg197 (**Figure 6B**).

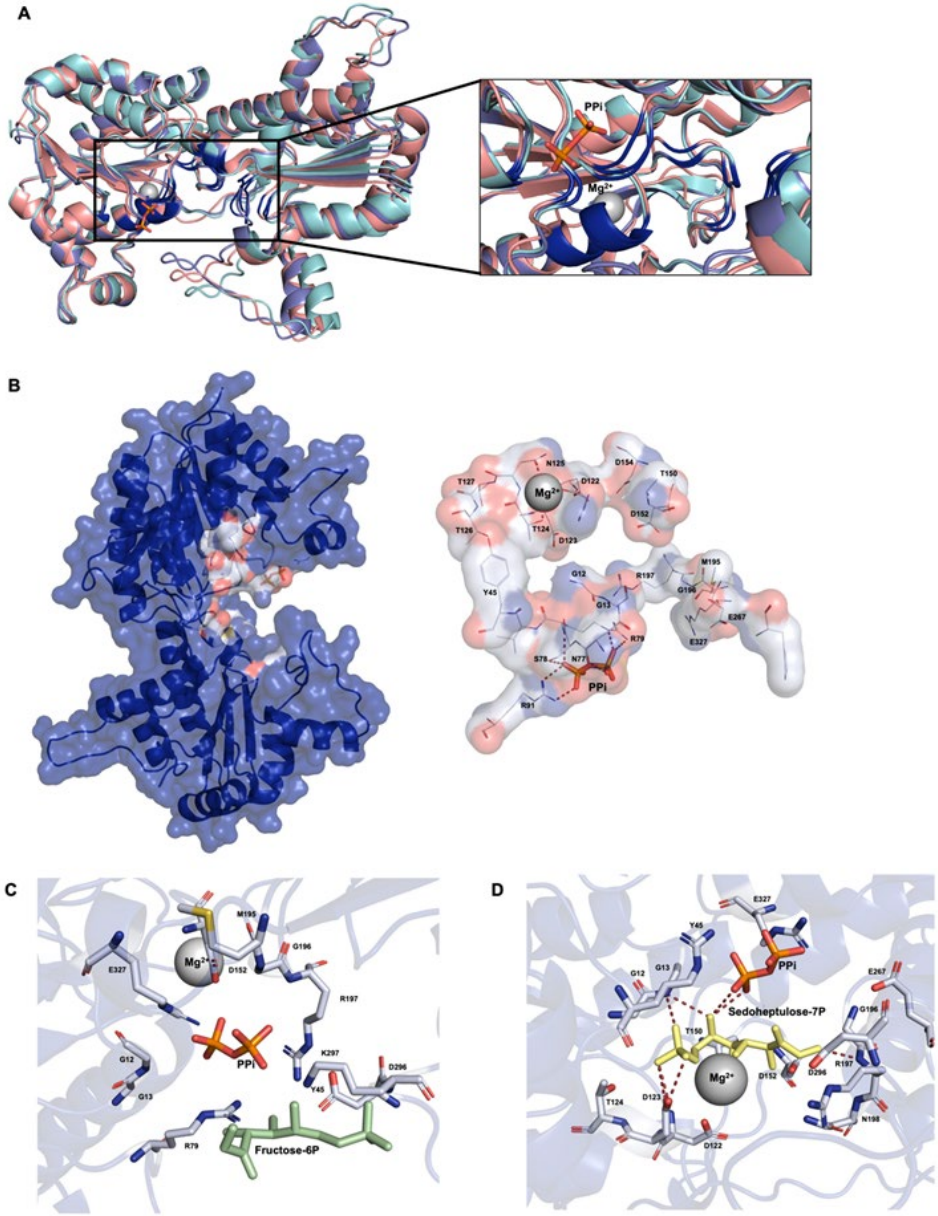

Figure 6. The PPi-PFK structure highlights the α/β domain and the catalytic domain interacting with the substrates. **(A)** The superimposed structures of PPi-PFK predicted using AlphaFold (violet), ESMFold (cyan) and RoseTTAFold (pink) shown in cartoon representation and the catalytic domain highlighted in blue color **(B)** The predicted model of the PPi-PFK protein of CLas gxpsy is shown in a cartoon representation in blue along with the catalytic domain shown in surface. The active site residues highlighting the PPi and Mg^{2+} bound region are shown in the right side. The PPi-PFK protein in complex with **(C)** F6P and **(D)** S7P after a simulation run of 100 ns.

Multiple sequence alignment of PPi-PFK sequence from CLas gxpsy shows variable similarity with Liberibacter spp., viz., CLaf (85%), CLam (74.5%), CLso (81.5%), and Lcr BT-1 (64.5%) and *E. coli* (22.5%), *Prevotella copri* (22%), *Clostridium thermosuccinogenes* (22.5%), and *Entamoeba histolytica* (21.6%). Besides the variable identity among the sequences, the alignment infers the conservation of the catalytic domain with less functional change (**Figure 7**). Based on the consensus sequence, *P.copri* is the closest homolog, and the non-synonymous substitutions were observed in *E. coli*, *C.thermosuccinogenes,* and *E.histolytica*.

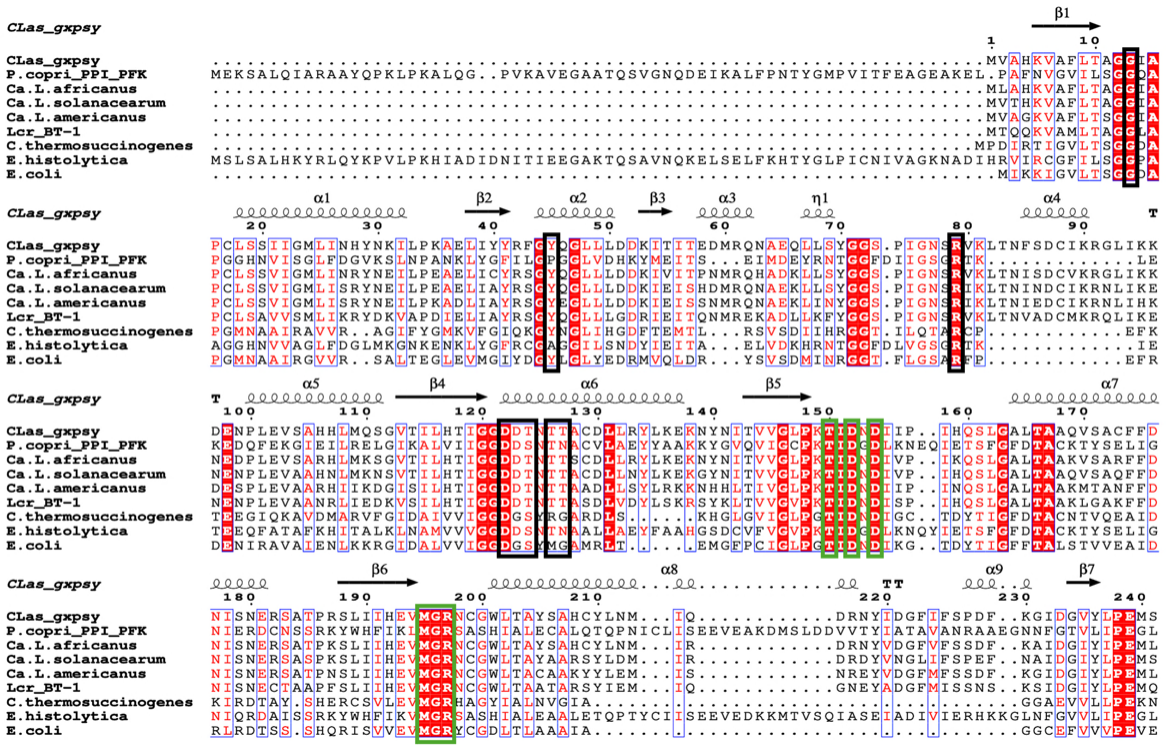

Figure 7. Multiple sequence alignment of PPi-PFK protein sequences from homologues species. The active sites are highlighted in black and green boxes and the substrate binding sites are shown in green boxes.

The rigid docking revealed that F6P and S7P formed a complex with the predicted structure with a binding energy of -5.6 and -5.7 kcal/mol, respectively. S7P and F6P showed good affinity in binding with the modeled structure. The molecular dynamics of the apo, PPi-PFK + Mg^{2+}, and PPi-PFK bound with F6P and S7P were studied to assess the structural stability of the complexes. The simulated structures of complexes were further analyzed in PyMOL [70]. The active site residues involved in the binding of S7P include Gly12, Gly13, Asp122, Asp123, Thr124, Thr150, Asp152, Gly196, Arg197, E267, and E327. The substrate S7P formed polar contact with three residues, Gly13, Asp123, Arg197, and PPi, in the binding pocket, making the interaction stronger (**Figure 6C**). The F6P formed interaction with the residues Tyr45, Arg79, and Arg197 in the binding pocket (**Figure 6D).** 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 [71]. The mutation studies revealed the key residue involved in the reaction is Asp127 of PFK-1 in *E. coli* which is equivalent to the Asp 152 of PPi-PFK in CLas. Also, the residue Arg171 of PFK-1 in *E. coli* plays a minor part in the catalysis, which aligns with Arg197 of PPi-PFK in CLas [69]. Structural studies have revealed the formation of a stable complex between PPi-PFK and its substrates, highlighting the involvement of key residues in substrate binding.

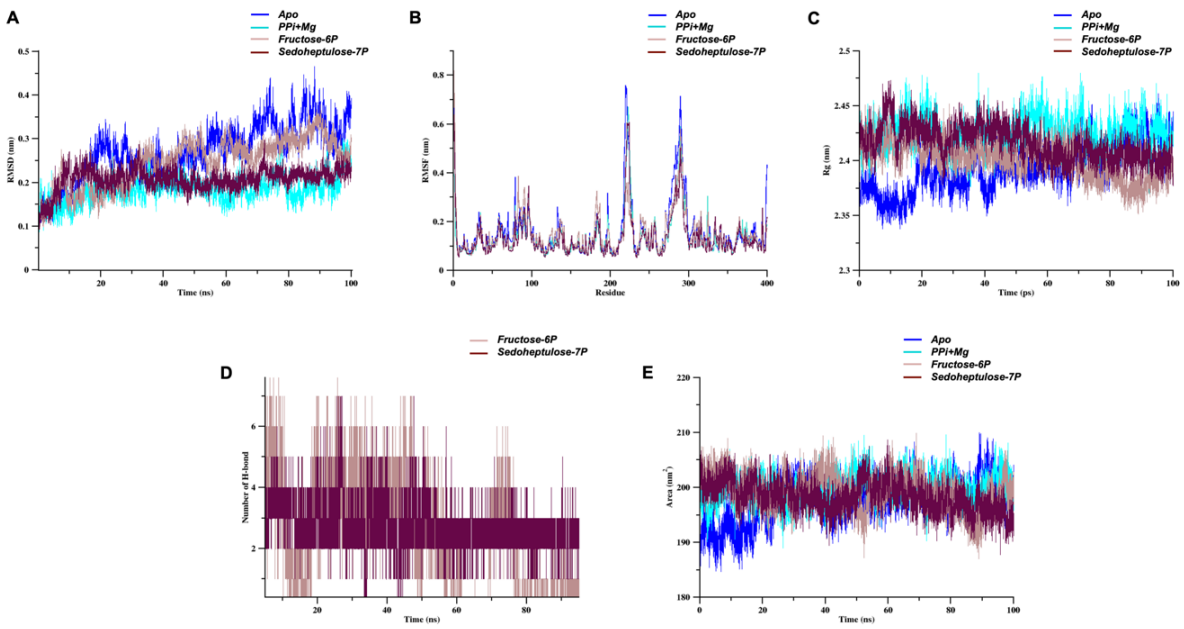

Figure 8. The structural stability analysis result for 100ns simulation run. The plots of (**A**) RMSD, (**B**) RMSF, (**C**) Rg, (**D**) number of H bond interaction and (**E**) SASA representing the stability of PPi-PFK and complexes throughout the trajectories.

The refined structures from molecular dynamics were analyzed using the obtained RMSD, RMSF, Rg, number of hydrogen bonds formed, and SASA values to understand the conformational stability. Each simulation run was for 100 ns. The RMSD graph depicted the average values of 0.26 nm, 0.18 nm, 0.24 nm, and 0.2 nm obtained for apo, PPi-PFK+Mg^{2+}, F6P, and S7P, respectively (**Figure 8A)**. RMSF is used to measure the fluctuation of residues of protein to form a stable complex with the ligands. The average RMSF values of 0.16 nm, 0.137 nm, 0.148, and 0.133 nm were attained for apo, PPi-PFK+Mg^{2+}, F6P, and S7P, respectively (**Figure 8B)**. Thus, indicating that the ligands formed a stable complex with the protein. Rg values help to understand the compactness of the protein. The average Rg values obtained for apo, PPi-PFK+Mg^{2+}, F6P, and S7P were 2.39 nm, 2.42 nm, 2.40 nm, and 2.41, respectively. Thus, implying the binding of the substrate didn't impact the packing of the

protein (**Figure 8C)**. 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 PPi-PFK were 2.46 and 2.4, respectively (**Figure 8D)**. The SASA evaluates the exposed residues of protein to the solvent molecules. The average values of SASA obtained for apo, PPi-PFK+Mg^{2+}, F6P, and S7P were 197.6 nm^2, 199.8 nm^2, 198.3 nm^2, and 199.2 nm^2, respectively, implying the stable formation of the complexes (**Figure 8E)**. The sequence and structural analysis revealed the possibility of the incorporation of the intermediate step to complete PPP (**Figure 1: Step 10**) and also helping in the conversion of F6P to FBP (**Figure 1: Step 13**).

3. Methodology

3.1 Sequence analysis

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 [72]. The genome of the CLas gxpsy 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 [73]. Sequences of prophage genome SC1 (NC_019549.1) and SC2 (NC_019550.1) were retrieved from NCBI. The genome sequences of Liberibacter spp. and their corresponding CDS sequences were fetched from NCBI with assembly "complete." The dataset of Liberibacter spp. was further divided based on the host from which each strain is isolated. Multiple genome sequence alignment using a Locally collinear blocks searching algorithm to identify lytic genes and lysogenic genes in Liberibacter spp. and Lcr BT-1 was carried out using Mauve software [64]. DBSCAN-SWA tool [65] was used to predict the regions of phage genes integrated in the bacterial genomes of Liberibacter spp. The results were analyzed and visualized using R Bioconductor package gggene [74]. The

protein sequence of pyrophosphate fructose 6-phosphate 1-phosphotransferase (PPi-PFK) (WP_015452326.1) was retrieved from NCBI and sequence subjected to the prediction of active site residues using NCBI conserved domain search [75]. The sequence similarity search was performed using BLASTP search against protein data bank database and analyzed catalytic domains of the structures. The multiple sequence alignment (MSA) was carried out for homologs selected based on the literature review using the MAFFT program [76] utilizing the FFT-NS-i iterative refinement method and BLOSUM62 as substitution matrix. The sequence similarity search and structural analysis of the Pyruvate dehydrogenase, E1 component alpha subunit pdhA (WSI_02605) from CLas gxpsy was carried out to annotate the protein.

3.2. Metabolic pathway analysis

The genome annotation of Lcr BT-1 and CLas gxpsy was carried out using the RAST annotation server [77], and the resulting genome sequence data was deployed to build a metabolic model using Kbase software [30]. The data of annotated transport genes in CLas gxpsy was obtained from Transport Classification Database (TCDB) [78].

3.3. Structure prediction

The three-dimensional structure of PPi-PFK protein found in CLas gxpsy was predicted using AlphaFold v2.1.1 [39], RoseTTAFold server [68] and ESMFold [67]. The predicted structures were compared with 1-phosphofructokinase (PDB ID: 1PFK) from *E. coli* to select the best structure model. The modeled proteins were further validated by checking stereochemistry using PROCHECK [79] in SAVES server v6.0. The validated model was modelled with cofactor pyrophosphate (PPi) and magnesium ion (Mg^{2+}) using Modeller 10.5 [80] and was refined using molecular dynamics simulation in GROMACS 2022.2 software [81] by

CHARMM36 forcefield [82]. The final model was visualized and analyzed in PyMOL software [70].

3.4. Molecular docking and structure refinement

The substrates fructose 6-phosphate (F6P) and sedoheptulose 7-phosphate (S7P) were retrieved from PubChem [83]. Both protein and ligand structures are pre-processed, and rigid docking was carried out using Autodock Vina in AMDock 1.5.2 graphical tool [84]. 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 were carried out using GROMACS software. The ligand topology and parameterization were performed using the CHARMM General Force Field (CGenFF) webserver [85]. In contrast, the CHARMM forcefield and TIP3P water module of GROMACS were utilized to generate protein topology. A unit cell of a dodecahedron shape with 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 NVT and NPT equilibration steps were carried out for 300K and 1 bar pressure, respectively. Finally, the system was subjected to a 100 ns simulation run. The parameters used to elucidate structural stability include Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Hydrogen bond formation, Solvent Accessible Surface Area (SASA), and Radius of gyration (Rg).

Conclusion

This study unravels the genomic streamlining, metabolic complexities, and genetic adaptations that enable CLas to survive as an obligate intracellular pathogen. Comparative metabolic analysis of CLas with Lcr BT-1 highlights the extensive reduction of essential metabolic

pathways in CLas, alongside the presence of unique or alternative routes that compensate for this loss. The CLas bacteria has lost genes encoded for pgi, PFK1, transaldolase, gloA which plays an essential part in the carbon metabolism and lacks PTS system to transport the sugar phosphate substrates directly. To circumvent metabolic constraints, CLas appears to regulate sugar substrate flow, potentially preventing the accumulation of toxic intermediates like methylglyoxal. The *in silico* structural studies revealed the activity of PPi-PFK protein, catalyzing two enzymatic reactions using fructose 6-phosphate and sedoheptulose 7-phosphate to sustain glycolysis. The presence of cytochrome bo3 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 retains acyl thioesterase to produce CoA and regulate fatty acid derivatives. Furthermore, the presence of type-1 lytic and type-2 lysogenic prophage genes suggests genetic adaptability. Further studies on critical metabolic pathways, virulence factors, and host interactions are essential to understanding the survival mechanisms of CLas within its host. Integrating *in vivo* plant-based and vector-based studies to understand CLas pathogenesis and *in vitro* approaches to establish an axenic culture will enable the biochemical and biophysical characterization of CLas. These insights will be crucial for developing effective disease management strategies and genetic manipulations.

Acknowledgement

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.

Author Contributions

SKP: Data curation, Methodology, Data analysis, Writing – original draft. JS & AKS: Conceptualization, Data analysis, Supervision, Writing – Review and Editing.

References

1 John P & Whatley FR (1975) Paracoccus denitrificans and the evolutionary origin of the mitochondrion. *Nature* **254**, 495–8.

2 Khademian M & Imlay JA (2021) How microbes evolved to tolerate oxygen. *Trends Microbiol* **29**, 428–440.

3 Martin WF & Sousa FL (2015) Early microbial evolution: The age of anaerobes. *Cold Spring Harb Perspect Biol* **8**, a018127.

4 Fridovich I (1989) Superoxide dismutases. An adaptation to a paramagnetic gas. *J Biol Chem* **264**, 7761–4.

5 Degli Esposti M, Mentel M, Martin W & Sousa FL (2019) Oxygen Reductases in Alphaproteobacterial Genomes: Physiological Evolution From Low to High Oxygen Environments. *Front Microbiol* **10**.

6 Bové JM (2006) Huanglongbing: a destructive, newly-emerging, century-old disease of citrus. . *J Plant pathol* **88**, 7–37.

7 Merfa M V., Pérez-López E, Naranjo E, Jain M, Gabriel DW & De La Fuente L (2019) Progress and Obstacles in Culturing ‘ *Candidatus* Liberibacter asiaticus’, the Bacterium Associated with Huanglongbing. *Phytopathology* **109**, 1092–1101.

8 Murray RG & Stackebrandt E (1995) Taxonomic note: implementation of the provisional status Candidatus for incompletely described procaryotes. *Int J Syst Bacteriol* **45**, 186–7.

9 Leonard MT, Fagen JR, Davis-Richardson AG, Davis MJ & Triplett EW (2012) Complete genome sequence of Liberibacter crescens BT-1. *Stand Genomic Sci* **7**, 271–283.

10 Jain M, Cai L, Fleites LA, Munoz-Bodnar A, Davis MJ & Gabriel DW (2019) *Liberibacter crescens* Is a Cultured Surrogate for Functional Genomics of Uncultured Pathogenic ‘ *Candidatus* Liberibacter’ spp. and Is Naturally Competent for Transformation. *Phytopathology* **109**, 1811–1819.

11 Cruz-Munoz M, Munoz-Beristain A, Petrone JR, Robinson MA & Triplett EW (2019) Growth parameters of Liberibacter crescens suggest ammonium and phosphate as essential molecules in the Liberibacter-plant host interface. *BMC Microbiol* **19**, 222.

12 Montano HG, Bertaccini A & Fiore N (2024) Phytoplasma-Associated Diseases in South America: Thirty Years of Research. *Microorganisms* **12**, 1311.

13 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.

14 Contaldo N, Bertaccini A, Paltrinieri S, Windsor H & Windsor GD (2012) Axenic culture of plant pathogenic phytoplasmas . *Phytopathol Mediterr* **51**, 607–617.

15 Sarhan MS, Hamza MA, Youssef HH, Patz S, Becker M, ElSawey H, Nemr R, Daanaa H-SA, Mourad EF, Morsi AT, Abdelfadeel MR, Abbas Mohamed T and Fayez M, Ruppel S & Hegazi NA (2019) Culturomics of the plant prokaryotic microbiome and the dawn of plant-based culture media - A review. *J Adv Res* **19**, 15–27.

16 Sechler A, Schuenzel EL, Cooke P, Donnua S, Thaveechai N, Postnikova E, Stone AL, Schneider WL, Damsteegt VD & Schaad NW (2009) Cultivation of ‘ *Candidatus* Liberibacter asiaticus’, ‘ *Ca.* L. africanus’, and ‘ *Ca.* L. americanus’ Associated with Huanglongbing. *Phytopathology* **99**, 480–486.

17 Ha PT, He R, Killiny N, Brown JK, Omsland A, Gang DR & Beyenal H (2019) Host-free biofilm culture of “Candidatus Liberibacter asiaticus,” the bacterium associated with Huanglongbing. *Biofilm* **1**, 100005.

18 Davis MJ, Mondal SN, Chen H, Rogers ME & Brlansky RH (2008) Co-cultivation of ‘ *Candidatus* Liberibacter asiaticus’ with Actinobacteria from Citrus with Huanglongbing. *Plant Dis* **92**, 1547–1550.

19 Rancati G, Moffat J, Typas A & Pavelka N (2018) Emerging and evolving concepts in gene essentiality. *Nat Rev Genet* **19**, 34–49.

20 Cai L, Jain M, Munoz-Bodnar A, Huguet-Tapia JC & Gabriel DW (2022) A synthetic ‘essentialome’ for axenic culturing of ‘Candidatus Liberibacter asiaticus.’ *BMC Res Notes* **15**, 125.

21 Wang N & Trivedi P (2013) Citrus huanglongbing: a newly relevant disease presents unprecedented challenges. *Phytopathology* **103**, 652–665.

22 Jain M, Munoz-Bodnar A & Gabriel DW (2017) Concomitant loss of the glyoxalase system and glycolysis makes the uncultured pathogen ``Candidatus Liberibacter asiaticus'' an energy scavenger. *Appl Environ Microbiol* **83**.

23 Vahling CM, Duan Y & Lin H (2010) Characterization of an ATP translocase identified in the destructive plant pathogen ``Candidatus Liberibacter asiaticus''. *J Bacteriol* **192**, 834–840.

24 Zhang S, Flores-Cruz Z, Zhou L, Kang B-H, Fleites LA, Gooch MD, Wulff NA, Davis MJ, Duan Y-P & Gabriel DW (2011) “Ca. Liberibacter asiaticus” carries an excision plasmid

prophage and a chromosomally integrated prophage that becomes lytic in plant infections. *Mol Plant Microbe Interact* **24**, 458–468.

25 Zheng Z, Bao M, Wu F, Van Horn C, Chen J & Deng X (2018) A Type 3 prophage of "Candidatus Liberibacter asiaticus" carrying a restriction-modification system. *Phytopathology* **108**, 454–461.

26 Fujiwara K, Iwanami T & Fujikawa T (2018) Alterations of Candidatus Liberibacter asiaticus-Associated Microbiota Decrease Survival of Ca. L. asiaticus in in vitro Assays. *Front Microbiol* **9**.

27 Moran NA (2002) Microbial minimalism: genome reduction in bacterial pathogens. *Cell* **108**, 583–586.

28 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**, 391–404.

29 Stincone A, Prigione A, Cramer T, Wamelink MMC, Campbell K, Cheung E, Olin-Sandoval V, Grüning N-M, Krüger A, Tauqeer Alam M, Keller MA, Breitenbach M, Brindle KM, Rabinowitz JD & Ralser M (2015) The return of metabolism: biochemistry and physiology of the pentose phosphate pathway. *Biol Rev Camb Philos Soc* **90**, 927–963.

30 Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, Joachimiak MP, Keegan KP, Kondo S, Kumar V, Land ML, Meyer F, Mills M, Novichkov PS, Oh T, Olsen GJ, Olson R, Parrello B, Pasternak

S, Pearson E, Poon SS, Price GA, Ramakrishnan S, Ranjan P, Ronald PC, Schatz MC, Seaver SMD, Shukla M, Sutormin RA, Syed MH, Thomason J, Tintle NL, Wang D, Xia F, Yoo H, Yoo S & Yu D (2018) KBase: The United States Department of Energy Systems Biology Knowledgebase. *Nat Biotechnol* **36**, 566–569.

31 Garschagen LS, Franke T & Deppenmeier U (2021) An alternative pentose phosphate pathway in human gut bacteria for the degradation of C5 sugars in dietary fibers. *FEBS J* **288**, 1839–1858.

32 Basen M & Kurrer SE (2021) A close look at pentose metabolism of gut bacteria. *FEBS J* **288**, 1804–1808.

33 Koendjbiharie JG, Hon S, Pabst M, Hooftman R, Stevenson DM, Cui J, Amador-Noguez D, Lynd LR, Olson DG & van Kranenburg R (2020) The pentose phosphate pathway of cellulolytic clostridia relies on 6-phosphofructokinase instead of transaldolase. *J Biol Chem* **295**, 1867–1878.

34 Susskind BM, Warren LG & Reeves RE (1982) A pathway for the interconversion of hexose and pentose in the parasitic amoeba Entamoeba histolytica. *Biochem J* **204**, 191–6.

35 Zuñiga C, Peacock B, Liang B, McCollum G, Irigoyen SC, Tec-Campos D, Marotz C, Weng N-C, Zepeda A, Vidalakis G, Mandadi KK & Borneman James and Zengler K (2020) Linking metabolic phenotypes to pathogenic traits among ``Candidatus Liberibacter asiaticus'' and its hosts. *NPJ Syst Biol Appl* **6**, 24.

36 Chakraborty S, Karmakar K & Chakravortty D (2014) Cells producing their own nemesis: Understanding methylglyoxal metabolism. *IUBMB Life* **66**, 667–678.

37 de Kok A, Hengeveld AF, Martin A & Westphal AH (1998) The pyruvate dehydrogenase multi-enzyme complex from Gram-negative bacteria. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* **1385**, 353–366.

38 Nelson DL & Cox MMichael (2005) *Lehninger principles of biochemistry. New York :W.H. Freeman* W.H. Freeman, New York.

39 Borkakoti N & Thornton JM (2023) AlphaFold2 protein structure prediction: Implications for drug discovery. *Curr Opin Struct Biol* **78**, 102526.

40 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**, 4210–4220.

41 D'mello R, Hill S & Poole RK (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 (N Y)* **142**, 755–763.

42 Al-Attar S, Yu Y, Pinkse M, Hoeser J, Friedrich T, Bald D & de Vries S (2016) Cytochrome bd Displays Significant Quinol Peroxidase Activity. *Sci Rep* **6**, 27631.

43 Lunak ZR & Noel KD (2015) A quinol oxidase, encoded by cyoABCD, is utilized to adapt to lower O2 concentrations in Rhizobium etli CFN42. *Microbiology (N Y)* **161**, 203–212.

44 Shahzad F, Tang L & Vashisth T (2023) Unraveling the mystery of canopy dieback caused by citrus disease Huanglongbing and its link to hypoxia stress. *Front Plant Sci* **14**, 1119530.

45 Hijaz F & Killiny N (2014) Collection and Chemical Composition of Phloem Sap from Citrus sinensis L. Osbeck (Sweet Orange). *PLoS One* **9**, e101830.

46 Molki B, Thi Ha P, Mohamed A, Killiny N, Gang DR, Omsland A & Beyenal H (2019) Physiochemical changes mediated by ``Candidatus Liberibacter asiaticus'' in Asian citrus psyllids. *Sci Rep* **9**, 16375.

47 Borisov VB, Gennis RB, Hemp J & Verkhovsky MI (2011) The cytochrome bd respiratory oxygen reductases. *Biochimica et Biophysica Acta (BBA) - Bioenergetics* **1807**, 1398–1413.

48 Batarseh TN, Batarseh SN, Morales-Cruz A & Gaut BS (2023) Comparative genomics of the Liberibacter genus reveals widespread diversity in genomic content and positive selection history. *Front Microbiol* **14**.

49 Pitino M, Armstrong CM & Duan Y (2017) Molecular mechanisms behind the accumulation of ATP and H2O2 in citrus plants in response to "Candidatus Liberibacter asiaticus" infection. *Hortic Res* **4**, 17040.

50 Tsu B V & Saier Jr MH (2015) The LysE superfamily of transport proteins involved in cell physiology and pathogenesis. *PLoS One* **10**, e0137184.

51 Welker S, Pierre M, Santiago JP, 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**, 379–390.

52 Jain M, Munoz-Bodnar A & Gabriel DW (2019) "Candidatus Liberibacter asiaticus" peroxiredoxin (LasBCP) suppresses oxylipin-mediated defense signaling in citrus. *J Plant Physiol* **236**, 61–65.

53 Bernardini C, Turner D, Wang C, Welker S, Achor D, Artiga YA, Turgeon R & Levy A (2022) *Candidatus* Liberibacter asiaticus accumulation in the phloem inhibits callose and reactive oxygen species. *Plant Physiol* **190**, 1090–1094.

54 Tarique KF, Abdul Rehman SA, 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**, 386–396.

55 Zakataeva NP (2021) Microbial 5'-nucleotidases: their characteristics, roles in cellular metabolism, and possible practical applications. *Appl Microbiol Biotechnol* **105**, 7661–7681.

56 Kumar P, Kesari P, Kokane S, Ghosh DK, Kumar P & Sharma AK (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**, 3450–3472.

57 Liu J, Zhang X, Deng S, Wang H & Zhao Y (2022) Thiamine Is Required for Virulence and Survival of Pseudomonas syringae pv. tomato DC3000 on Tomatoes. *Front Microbiol* **13**, 903258.

58 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.

59 Liu K, He J, Guan Z, Zhong M, Pang R & Han Q (2020) Transcriptomic and metabolomic analyses of Diaphorina citri Kuwayama infected and non-infected with Candidatus Liberibacter asiaticus. *Front Physiol* **11**, 630037.

60 Brüssow H, Canchaya C & Hardt W-D (2004) Phages and the Evolution of Bacterial Pathogens: from Genomic Rearrangements to Lysogenic Conversion. *Microbiology and Molecular Biology Reviews* **68**, 560–602.

61 Carey JN, Mettert EL, Fishman-Engel DR, Roggiani M, Kiley PJ & Goulian M (2019) Phage integration alters the respiratory strategy of its host. *Elife* **8**.

62 Jain M, Fleites LA & Gabriel DW (2015) Prophage-Encoded Peroxidase in ‘ *Candidatus* Liberibacter asiaticus’ Is a Secreted Effector That Suppresses Plant Defenses. *Molecular Plant-Microbe Interactions®* **28**, 1330–1337.

63 Hao G, Ammar D, Duan Y & Stover E (2019) Transgenic citrus plants expressing a ‘Candidatus Liberibacter asiaticus’ prophage protein LasP235 display Huanglongbing-like symptoms. *Agri Gene* **12**, 100085.

64 Darling ACE, Mau B, Blattner FR & Perna NT (2004) Mauve: Multiple Alignment of Conserved Genomic Sequence With Rearrangements. *Genome Res* **14**, 1394–1403.

65 Gan R, Zhou F, Si Y, Yang H, Chen C, Ren C, Wu J & Zhang F (2022) DBSCAN-SWA: An Integrated Tool for Rapid Prophage Detection and Annotation. *Front Genet* **13**.

66 Zheng Z, Bao M, Wu F & Chen Jianchi and Deng X (2016) Predominance of single prophage carrying a CRISPR/cas system in ``Candidatus Liberibacter asiaticus'' strains in southern China. *PLoS One* **11**, e0146422.

67 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**.

68 Krishna R, Wang J, Ahern W, Sturmfels P, Venkatesh P, Kalvet I, Lee GR, Morey-Burrows FS, Anishchenko I, Humphreys IR, McHugh R, Vafeados D, Li X, Sutherland GA, Hitchcock A, Neil Hunter C, Kang A, Brackenbrough E, Bera AK, Baek M, DiMaio F & Baker D (2024) Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *Science (1979)* **384**.

69 Shirakihara Y & Evans PR (1988) Crystal structure of the complex of phosphofructokinase from Escherichia coli with its reaction products. *J Mol Biol* **204**, 973–94.

70 Seeliger D & de Groot BL (2010) Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J Comput Aided Mol Des* **24**, 417–422.

71 Moore SA, Ronimus RS, Roberson RS & Morgan HW (2002) The structure of a pyrophosphate-dependent phosphofructokinase from the Lyme disease spirochete Borrelia burgdorferi. *Structure* **10**, 659–671.

72 Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J & Sayers EW (2012) GenBank. *Nucleic Acids Res* **41**, D36–D42.

73 Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* **215**, 403–410.

74 Reimers M & Carey VJ (2006) Bioconductor: an open source framework for bioinformatics and computational biology. *Methods Enzymol* **411**, 119–134.

75 Marchler-Bauer A, Anderson JB, Derbyshire Myra K and DeWeese-Scott C, Gonzales NR, Gwadz Marc and Hao L, He S, Hurwitz DI, Jackson JD, Ke Z, Krylov D, Lanczycki CJ, Liebert CA, Liu C, Lu Fu and Lu S, Marchler GH, Mullokandov M, Song JS, Thanki N, Yamashita RA, Yin JJ, Zhang D & Bryant SH (2007) CDD: a conserved domain database for interactive domain family analysis. *Nucleic Acids Res* **35**, D237–40.

76 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**, 3059–3066.

77 Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A & Zagnitko O (2008) The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics* **9**, 75.

78 Saier MH, Reddy VS, Moreno-Hagelsieb G, Hendargo KJ, Zhang Y, Iddamsetty V, Lam KJK, Tian N, Russum Steven and Wang J & Medrano-Soto A (2021) The Transporter Classification Database (TCDB): 2021 update. *Nucleic Acids Res* **49**, D461–D467.

79 Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R & Thornton JM (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J Biomol NMR* **8**, 477–86.

80 Webb B & Sali A (2014) Protein structure modeling with MODELLER. *Methods Mol Biol* **1137**, 1–15.

81 Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE & Berendsen HJC (2005) GROMACS: fast, flexible, and free. *J Comput Chem* **26**, 1701–1718.

82 Huang J & MacKerell AD (2013) CHARMM36 all-atom additive protein force field: Validation based on comparison to NMR data. *J Comput Chem* **34**, 2135–2145.

83 Li Q, Cheng T, Wang Y & Bryant SH (2010) PubChem as a public resource for drug discovery. *Drug Discov Today* **15**, 1052–1057.

84 Valdés-Tresanco MS, Valdés-Tresanco ME, Valiente PA & Moreno E (2020) AMDock: a versatile graphical tool for assisting molecular docking with Autodock Vina and Autodock4. *Biol Direct* **15**, 12.

85 Vanommeslaeghe K, Raman EP & MacKerell Jr AD (2012) Automation of the CHARMM General Force Field (CGenFF) II: assignment of bonded parameters and partial atomic charges. *J Chem Inf Model* **52**, 3155–3168.