CINDY LEE CAJACHAGUA PUCUHUARANGA

Comportamento estrutural de PstS e PhoX de *Xanthomonas citri* subsp. *citri* na captação de fosfato e sua relação com proteínas quimiotáticas

Dissertação apresentada ao Programa de Pós-graduação em Microbiologia do Instituto de Ciências Biomédicas da Universidade de São Paulo, para obtenção do título de Mestre em Ciências

São Paulo
2021

CINDY LEE CAJACHAGUA PUCUHUARANGA

Structural behavior of PstS and PhoX from *Xanthomonas citri* subsp. *citri* in the uptake of phosphate and their relation with chemotactic proteins

Dissertation presented to the post-graduate program in Microbiology at the Institute of Biomedical Sciences at the University of São Paulo for the Master of Science degree

São Paulo
2021

CINDY LEE CAJACHAGUA PUCUHUARANGA

Comportamento estrutural de PstS e PhoX de *Xanthomonas citri* subsp. *citri* na captação de fosfato e sua relação com proteínas quimiotáticas

Dissertação apresentada ao Programa de Pós- Graduação em Microbiologia do Instituto de Ciências Biomédicas da Universidade de São Paulo, para obtenção do Título de Mestre em Ciências

Área de concentração: Microbiologia

Orientadora: Prof. Dra. Andrea Balan Fernandes

Versão original

São Paulo
2021

CINDY LEE CAJACHAGUA PUCUHUARANGA

Structural behavior of PstS and PhoX from *Xanthomonas citri* subsp. *citri* in the uptake of phosphate and their relation with chemotactic proteins

Dissertation presented to the graduated program in Microbiology at the Institute of Biomedical Sciences at the University of São Paulo for the Master of Science degree

Concentration area: Microbiology

Supervisor: Prof. Dra. Andrea Balan Fernandes

Original version

São Paulo
2021

UNIVERSIDADE DE SÃO PAULO INSTITUTO DE CIÊNCIAS BIOMÉDICAS

Candidato(a):   Cindy Lee Cajachagua Pucuhuaranga

Título da dissertação: Structural behavior of PstS and PhoX from *Xanthomonas citri* subsp. *citri* in the uptake of phosphate and their relation with chemotactic proteins

Orientador: Profa. Dra. Andrea Balan Fernandes

A comissão julgadora dos trabalhos de defesa da Dissertação de Mestrado, em sessão pública realizada ..../........./ ...., considerou

( ) Aprovado(a)   ( ) Reprovado(a)

Examinador(a):   Assinatura: ....
                   Nome: ....
                   Instituição: ....

Examinador(a):   Assinatura: ....
                   Nome: ....
                   Instituição: ....

Examinador(a):   Assinatura: ....
                   Nome: ....
                   Instituição: ....

Presidente: Assinatura: ....
                   Nome: ....
                   Instituição: ....
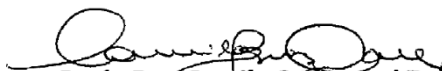
ICB USP

CERTIFICADO DE ISENÇÃO

Certificamos que o Protocolo CEP-ICB nº **928/2018** referente ao projeto intitulado: *"Estudos estruturais e funcionais comparativos entre Phox e PsTS de Xanthomonas axonopodis pv. Citri"* sob a responsabilidade de *Cindy Lee Cajachagua Pucuhuaranga* e orientação do(a) Prof.(a) Dr.(a) *Andrea Balan Fernandes*, do Departamento de Microbiologia, foi analisado pela **CEUA** - Comissão de Ética no Uso de Animais e pelo **CEPSH** – Comitê de Ética em Pesquisa com Seres Humanos, tendo sido deliberado que o referido projeto não utilizará animais que estejam sob a égide da Lei nº 11.794, de 8 de outubro de 2008, nem envolverá procedimentos regulados pela Resolução CONEP nº 466/2012.

São Paulo, 26 de fevereiro de 2018.

Profa. Dra. **Luciane Valéria Sita**
Coordenadora CEUA ICB/USP

Profa. Dra. **Camila Squarzoni Dale**
Coordenadora CEPSH ICB/USP

EXPEDIÇÃO
DATA 9 / 3 / 18
REL. Nº 229

À minha irmã Ketty Liz Cajachagua, que sempre me apoiou e guiou com paciência e amor. Quem me ensinou que embora a vida tenha milhares de obstáculos, vale a pena viver e lutar pelos nossos sonhos. Porque uma pessoa pode ser feliz, e o primeiro passo para alcançar a felicidade é o autoconhecimento.

A mi hermana Ketty Liz Cajachagua, que siempre me ha apoyado y guiado con paciencia y cariño. Quien me ha enseñado que, aunque la vida tenga miles de obstáculos, vale la pena vivir y luchar por nuestros sueños. Porque una persona si puede ser feliz, y el primer paso para alcanzar la felicidad es el autoconocimiento.

ACKNOWLEDGMENT

Agradeço aos meus pais Teresa Pucuhuaranga e Cirilo Cajachagua que sempre me apoiaram para alcançar todos os objetivos que me propus e que me ensinaram que na vida o mais importante é ser uma pessoa boa e desejar o melhor para todos, porque a vida é muito curta para desperdiçá-la com sentimentos ruins.

À minha orientadora Andrea Balan Fernandes pela oportunidade, paciência, compreensão, amizade, pelos conselhos, pelo amor pela ciência que me ajudou a não desistir. Ao Dr. Finan Turlough que me ensinou que não se trata de ter um grande número de resultados, trata-se de saber interpretá-los bem, e porque ele me apoiou durante os tempos difíceis da pandemia com suas palavras de encorajamento e sua amizade.

À minha família em São Paulo:

Aos meus amigos do Laboratório de Biologia Estrutural; Fernanda, Sindy, Lilia, Marcelo, Brenda, Danilo, Laura, Catharina, Sair e Pamela, que me fizeram sentir em casa todo o tempo que estive no Brasil, e por compartilharem comigo horas de estudo, choro e risos. Ao Lucas, Stephany e Victor, que me deram sua amizade desde o primeiro segundo que me conheceram e que me deram apoio emocional e acadêmico durante meus últimos meses no Brasil.

À Tania e Amzy pela companhia durante as noites no laboratório, por estar comigo nos momentos bons e ruins, por me incentivar academicamente e por me ajudar a ser uma pessoa melhor.

À Maryam, Leah, Jason e Barney pelos ensinamentos de vida, as piadas e as experiências no laboratório.

Ao Edgar, Ale, Guillermo, Henry e Diana pelos belos momentos compartilhados, e por suportarem meu mau humor.

"It is only with the heart that one can see rightly; what is essential is invisible to the eye. The most beautiful things in the world cannot be seen or touched, they are felt with the heart."

The Little Prince
Antoine de Saint-Exupéry (1943)

RESUMO

Xanthomonas citri subsp. *citri* é uma Gammaproteobactéria causadora do cancro cítrico em plantas cítricas, afetando a qualidade dos frutos e causando prejuízos econômicos. Para o controle da doença, esforços têm sido feitos no sentido de compreender sua fisiologia e mecanismos patogênicos. O fosfato é essencial para as bactérias em geral, como para o gênero *Xanthomonas* uma vez que desempenha papéis importantes no metabolismo celular. Durante as condições de privação de fosfato, um dos sistemas mais importantes responsáveis pela captação deste ânion é o sistema Pst (PstSCAB-PhoU), no qual PstS é a proteína ligadora de fosfato com alta afinidade (*phosphate-binding protein*, PBP), localizada no periplasma. No caso de *X. citri*, duas PBPs foram anotadas no seu genoma, PstS e seu parálogo denominado PhoX. Na presente pesquisa, essas duas proteínas, foram estudadas em um nível estrutural por métodos biofísicos. Técnicas como Dicroísmo Circular (CD), Espalhamento de raios-X a baixo ângulo (SAXS), fluorescência dos triptofanos (TSF) e dinâmica molecular (MD) mostraram que PstS e PhoX, apesar de conservadas estrutural e funcionalmente, apresentam diferenças relevantes no seu comportamento em solução, estabilidade e flexibilidade. A afinidade de PhoX para o fosfato foi determinada pela primeira vez e apresenta um Kd de 6.1 µM. Adicionalmente, análises de bioinformática mostraram que PhoX e PstS conservam resíduos carregados positivamente com proteínas ligadoras de ânions que podem ter papel importante no processo de transporte do ânion. A relação entre essas duas PBPs e a quimiotaxia em *X. citri* não foi estabelecida, mas a partir da obtenção de cepas mutantes com deleção única dos genes *pst*S e *pho*X poderá ser estudada futuramente na presença e ausência de fosfato.

Palavras-chave: Dinâmica molecular**.** Proteínas ligadoras periplasmáticas. Captação de fosfato. Afinidade por fosfato. Sistema Pst.

ABSTRACT

CAJACHAGUA, C. **Structural behavior of PstS and PhoX from *Xanthomonas citri* subsp. *citri* in the uptake of phosphate and their relation with chemotactic proteins.** 2021. 87 p. Master thesis (Microbiology) – Institute of Biomedical Sciences, University of Sao Paulo, SP, 2021.

Xanthomonas citri subsp. *citri* is a Gammaproteobacteria that causes the canker disease in citrus plants affecting the quality of fruits and causing economic losses. To control this disease, efforts have been made to understand its physiology and pathogenic mechanisms. Phosphate is essential for bacteria as well as for *Xanthomonas* genus once it plays important role in cell metabolism. During conditions of phosphate starvation, one of the most important systems responsible for the anion uptake is Pst system (PstSCAB-PhoU), where PstS is the protein that bind phosphate molecules with high affinity (*phosphate-binding protein*, PBP), and is located on the periplasm. In the case of *X. citri*, two PBPs were annotated, PstS and its paralog named PhoX. In the present research, these proteins were studied at a structural level using biophysics methods. Techniques such as Circular dichroism (CD), Small-angle X-ray scattering (SAXS), Tryptophan fluorescence spectroscopy (TSF) and molecular dynamic simulation (MD) showed that, despite being structurally and functionally conserved, PstS and PhoX present relevant differences, such as behavior in solution, stability, and flexibility. The affinity of PhoX for phosphate was determined for the first time with Kd value of 6.1 μM. Furthermore, bioinformatics analyzes showed that PhoX and PstS conserve residues positively charged between periplasmic binding proteins that may play an important role in the transport process of anions. The relationship between these two PBPs and chemotaxis in *X. citri* was not stablished but from the obtaining of mutant strains with single deletion of genes *pst*S and *pho*X this can be further studied in the presence and absence of phosphate.

Key words: Molecular dynamics. Periplasmic-binding protein. Phosphate uptake. Phosphate affinity. Pst system.

LISTS OF FIGURES

LISTS OF TABLES

LIST OF ABREVIATIONS AND ACRONYMS

CC - Citrus canker

CD - Circular dichroism

GROMACS - Groningen machine for chemical simulation

HAMP - Histidine kinase, adenyl cyclase, methyl-accepting chemotaxis protein and phosphatase domain

HK - Histidine Kinase

IMAC - Immobilized nickel-affinity chromatography

K_D - The equilibrium binding constant

LBD - Ligand binding domains

MCPs - Methyl-accepting chemotaxis proteins

MWCO - Molecular weight cut-off

NBD – Nucleotide Binding Domains

PAS - Per-Arnt-Sim domain

PBPs - Phosphate Binding Proteins

PeBPs - Periplasmic Binding Proteins

PCR - Polymerase chain reaction

Pi - Inorganic phosphate

PVDF - Polyvinylidene difluoride membrane

RMSD - Root-mean-square deviation

RMSF - Root-mean-square fluctuation

RR - Response Regulator

SAXS - Small-angle X-ray scattering

SEC - Size-exclusion chromatography

SBDs - Substrate binding domains

SBPs - Substrate Binding Proteins

SDS-PAGE - Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis

TCS – Two Component System

TFS - Tryptophan fluorescence spectroscopy

TMDs - Transmembrane domains

Trp - Tryptophan

VMD - Virtual molecular dynamics

SUMMARY

1 INTRODUCTION

1.1 *Xanthomonas* genus

Xanthomonas is a Gram-negative bacterial genus belonging to the Gammaproteobacteria class and widely studied because encompass a large group of plant-associated bacteria that usually cause diseases in several agricultural commodities. According to the origin of the word, *Xanthomonas* comes from two Greek words; *xanthos*, meaning 'yellow', and *monas*, meaning 'entity', these words define a typical characteristic of most of the species of the *Xanthomonas* genus, which is the production of a unique yellow pigment called xanthomonadin. Additionally, several species have a mucoid texture **(Figure 1A)** and produce xanthan gum (SHARMA *et al.*, 2014; FONSECA *et al.*, 2019; TANG *et al.*, 2020). Morphologically *Xanthomonas* cells are short and straight rods, monotrichous with a polar flagellum, and have a size in the range of 0.4-1.0 x 1.2-3.0 μm **(Figure 1B)**. The optimum growth temperature is in the range 25-30°C (SHARMA *et al.*, 2014; AN *et al.,* 2020).
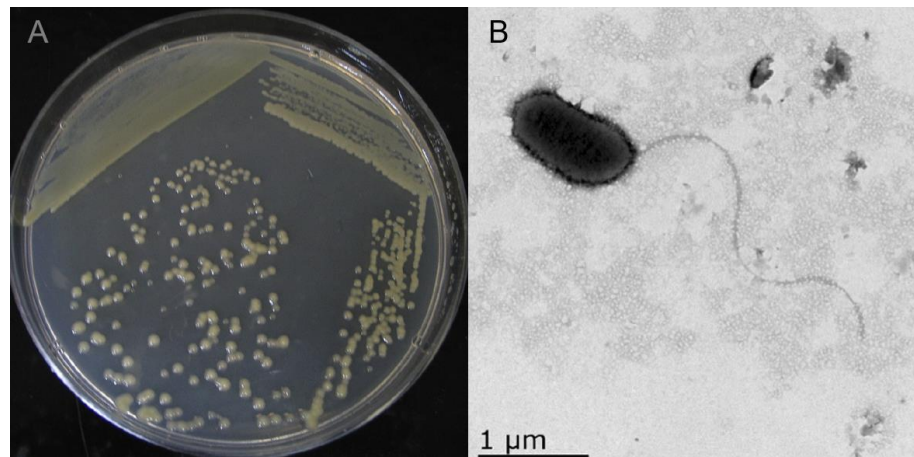
Figure 1. *Xanthomonas* **colonies and morphology. A.** Colonies of *Xanthomonas spp.* are yellow and mucoid on a nutrient agar plate. **B.** Transmission electron microscopy (TEM) image of *Xanthomonas oryzae* pv. *oryzae* detecting the flagella, shape, and size of the bacteria (Source: SHARMA *et al.*, 2014; LUO *et al.*, 2020).

The genus currently comprises more than 35 species and it is subdivided into pathovars. A 'Pathovar' (pv.) is defined as a group of strains that cause the same disease on the same host range, in this respect, pathovars have no taxonomical standing but it is

practical for plant pathologists and regulation purposes (TIMILSINA *et al.*, 2020; CATARA *et al.*, 2021). At least 125 pathovars have been described within the species of this genus. It was identified more than 400 different plant hosts for *Xanthomonas* species (spp.), such as rice (*X. oryzae*), cabbage (*Brassica* spp.), potato (*X. campestris*, *Solanum* spp.), orange (*X. vesicatoria*) and *Citrus* spp. (*X. citri*). The wide range of plant host is the reason why this genus causes huge economic losses to the agricultural sector (FONSECA *et al.*, 2019). The plant hosts include monocots and dicots, and pathogenic species of *Xanthomonas* can invade either the xylem elements of the vascular system (forming the group of vascular pathogens) or the intercellular spaces of the mesophyll parenchyma tissue (forming the group of mesophyllic pathogens), producing localized infections or systemic infection in case the pathogen can colonize both mesophyllic and vascular tissue (AN *et al.*, 2020; TIMILSINA *et al.*, 2020) **(Figure 2A).**

The transmission route of *Xanthomonas* spp. can be via contaminated seeds, spread to healthy plants by agricultural practices in contaminated soil (could be weeds or plant debris infected that constitute a potential source of inoculum), by pruning and misting of seed beds, by rainwater, formation of aerosols that can carry bacteria from adjacent infected fields and, possibly, by the fly of insects from one place to another (RYAN *et al.,* 2011; An *et al.,* 2020).

Pathogenic bacteria initially grow on leaf surfaces (epiphytic stage) and then infect the host (endophytic stage). The infection can be via stomates to colonize the intercellular spaces of the mesophyll parenchyma, or via hydathodes (water pores at the leaf margin) or wounds to spread systemically through the vascular system (SHARMA *et al.*, 2014; RYAN *et al.,* 2011). The type of affected physiological function depends on the cells and tissues of the host plant. Thus, the infection of xylem vessels interferes with the translocation of water, leading to vascular wilts and cankers, whereas infection of foliage interferes with the photosynthetic process and generate leaf spots, blights, and pustules (SHARMA *et al.*, 2014). Many *Xanthomonas* spp. can survive in the soil in association with dead plant material to form a source of inoculum for future crops, which can also become infected through wounds in the roots (RYAN *et al.,* 2011) **(Figure 2B).** Due to the genetic diversity of the pathovars, with varying degrees of aggressiveness and virulence, control and eradication efforts are complicated (FONSECA *et al.,* 2019).
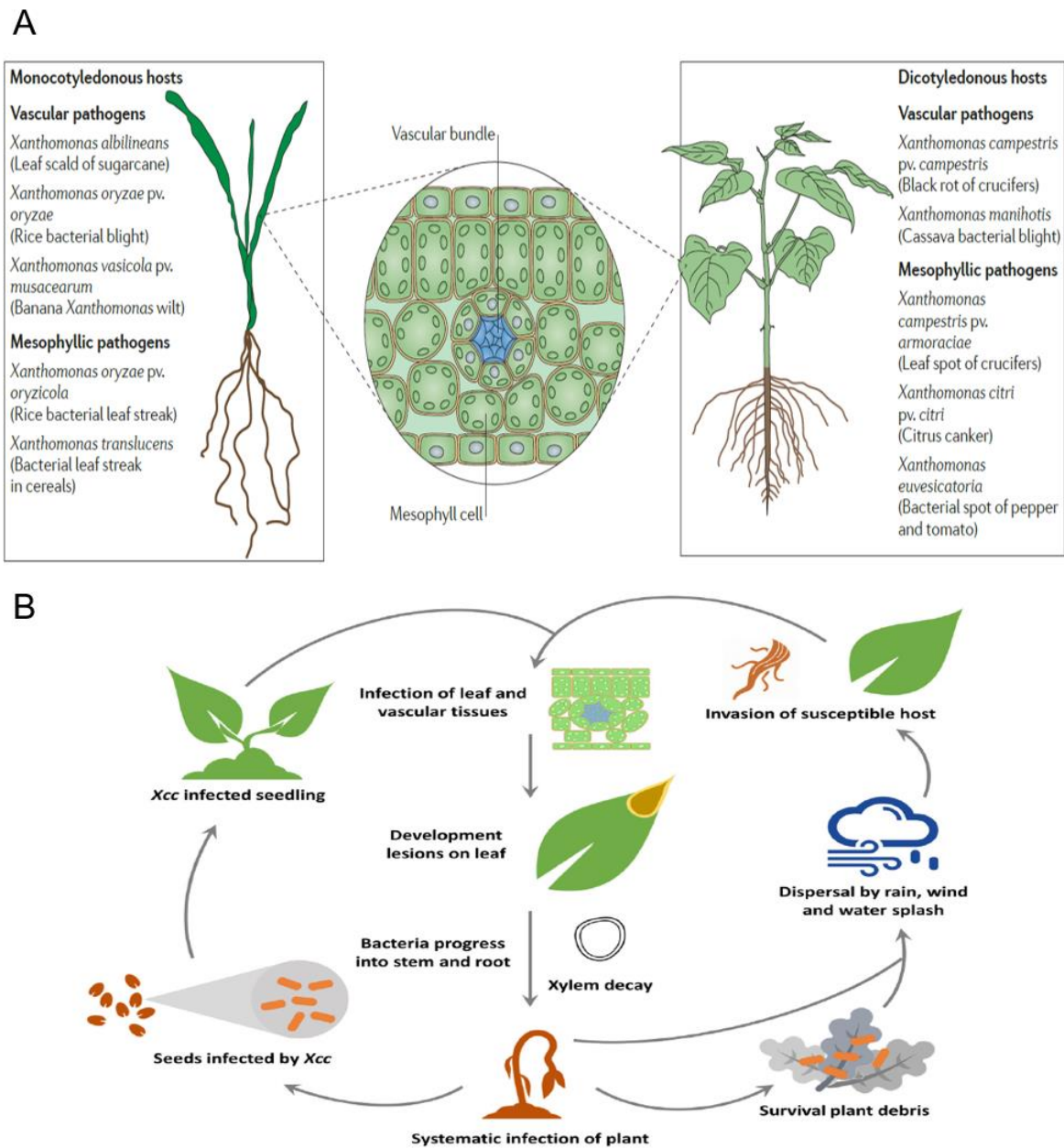
Figure 2. *Xanthomonas* **host invasion and infection A.** Plant hosts of pathogenic *Xanthomonas* spp. include monocots and dicots, and can infect the xylem elements of the vascular system or the intercellular spaces of the mesophyll tissue dividing the pathogens in two groups: vascular or mesophyllic pathogens, respectively. However, some bacteria are capable of colonizing both mesophyllic and vascular tissue like *Xanthomonas manihotis*, a cassava pathogen. **B.** Model illustrating the life cycle and disease symptoms of black rot pathogen *Xanthomonas campestris* pv. *campestris* (*Xcc*). Like most Xanthomonads, *Xcc* can survive in plant debris or free soil. In the case of *Xcc,* it also has the ability to colonise plant seeds (which represents a major route of disease transmission) and can also infect healthy plants by environmental and mechanical means. After germination of colonised seeds, the seedling becomes infected and the disease may manifest as shrivelling and the blackening of the margins of the seedling. *Xcc* may also invade mature plants via hydathodes, although leaf damage caused by insects and the root system also serve as portals of entry. These entry points usually provide a direct path to the plant vascular system leading to systemic host infection. Necrotic V-shaped lesions extending from the leaf margins manifest as the infection develops (Source: RYAN *et al.* 2011; An *et al.* 2020).

1.1.1 Citrus Canker and Xanthomonas citri subsp. citri

Citrus canker (CC) is a disease caused primarily by *Xanthomonas citri subsp. citri* (*X. citri*) in reference to the etiological agent of citrus canker type A, the most widespread type of CC (PATANÉ *et al.,* 2019; CAICEDO AND VILLAMIZAR *et al.*, 2021). This disease affects citrus production in several tropical and subtropical areas of the world (BOCK *et al.*, 2011), including Brazil, which is the largest producer of orange fruit worldwide (FAOSTAT, 2019), leading to important economic losses. The infection cycle of *X. citri* is as described for the *Xanthomonas* genus involving the epiphytic stage and the endophytic stage. Additionally, there is evidence that *X. citri*, as mesophyllic pathogen, can spread the infection through stomata. The bacteria gain entrance through stomata and colonize the apoplast, then produce the characteristic canker on the leaves and finally the leaf epidermis is disrupted as a result of cell hyperplasia induced by the pathogen allowing the released of bacteria that are dispersed by wind and rain initiating a new cycle **(Figure 3)** (GOTTIG *et al.*, 2010). CC is not systemic; it causes local lesions only. The pathogen causes distinctive necrotic raised lesions on leaves, branches, and fruits. Lesions that develop on the fruits make them unfit for commercializing in fresh and international markets **(Figure 4)** (GOTTIG *et al.,* 2010; BOCK *et al.*, 2011; RICHARD *et al.,* 2017). Severe infections can cause defoliation, blemished fruit, premature fruit drop, twig dieback, and general tree decline (SCHUBERT *et al.*, 2001).
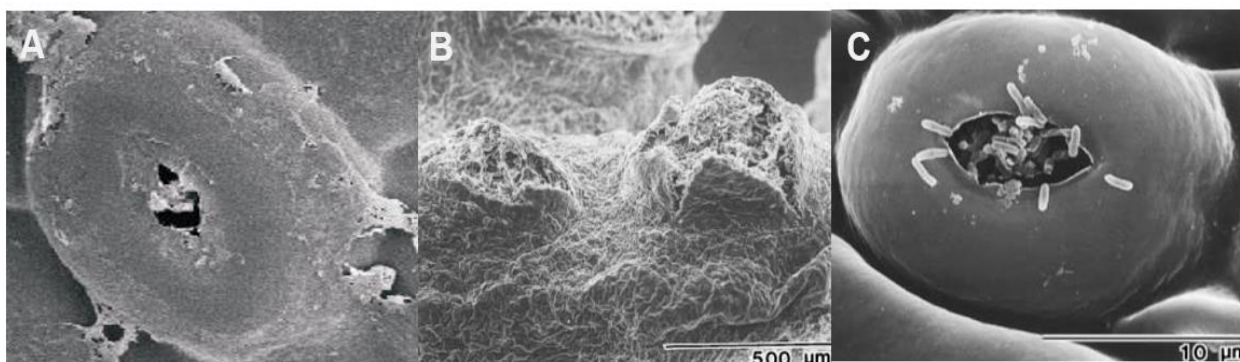
Figure 3. Scanning Electron Microscopy (SEM) imagens of *X. citri* **spread through stomata. A.** Image of the stomata entry which is the mainly entering of plants for *X. citri.* **B.** Image of of infected mesophyll tissues through the epidermis to form a pustule 30 days after inoculation. **C.** Image of bacterial egress from a stomatal opening allowing the disperse by wind and rain (Source: GRAHAM *et al.*, 2004; GOTTIG *et al.*, 2010).

Figure 4. Different types of lesions produced on *Citrus* spp. **plants by** *X. citri*. **A.** Yellow halo and raised lesions on leaf surface. **B.** Lesions on twig. **C.** Lesions on fruit. (Source: GOTTWALD *et al.*; 2002; SHARMA *et al.*, 2009; AN *et al.*, 2020).

Until now, measures to control the CC disease involve two different situations: locations with newly established of outbreak and locations where CC has become endemic (GOTTWALD *et al.*, 2002; DAS, 2003). In the first situation, programs of eradication are successful. Once *X. citri* is introduced into an area, elimination and destruction of infected and exposed trees is accepted to contain the disease and stop further spread. For instance, in Brazil if plantings have 0.5% infection or less, all trees within 30 m of infected trees are removed, whereas if infection is greater than 0.5%, the entire block is removed (GOTTWALD *et al.*, 2002). In locations with endemic CC, the use of copper-based bactericides helps in decreasing the disease severity but is not able to eradicate the disease (DHAKAL *et al.*, 2009). However, its extensive use can generate copper-resistant (CuR) strains (NAYEM *et al.*, 2017; GOCHEZ *et al.*, 2018). Moreover, other bactericides, including some antibiotics, are not as effective as copper because they lack sufficient residual activity to protect leaf and fruit surfaces for extended periods (HASABI *et al.*, 2014; MENDOÇA *et al.*, 2017). The problem with these chemical methods is the increase of cost production and environmental pollution (NAYEM *et al.*, 2017).

The genome of *X. citri* has been completely sequenced and gives insights into the importance of genes and gene clusters governing strategic mechanisms of pathogenesis. To develop efficient strategies, sustainable and economically viable to control the citrus canker disease, it is necessary to know *X. citri* physiology and its pathogenic mechanisms (GOTTIG *et al.*, 2010). During pathogen-host interaction some mechanisms are crucial such as motility (swimming and swarming), adhesion to the host cell, colonization, dissemination, and strategies to prolong their survival in the host like the biofilm formation (GOTTIG *et al.*, 2010; MARCIO *et al.*, 2014). Uptake of phosphate and chemotaxis are

two important events required for the initial colonization. The availability and uptake of inorganic phosphate (Pi) interferes directly in biofilm formation and actives complex processes of bacterial virulence (LAMARCHE *et al.*, 2008). On the other hand, the chemotactic processes, involves the motility toward a gradient concentration of root exudates or eliminated exudates through natural openings and lesions in plant tissues (MARCIO *et al.*, 2014) **(Figure 5)**.
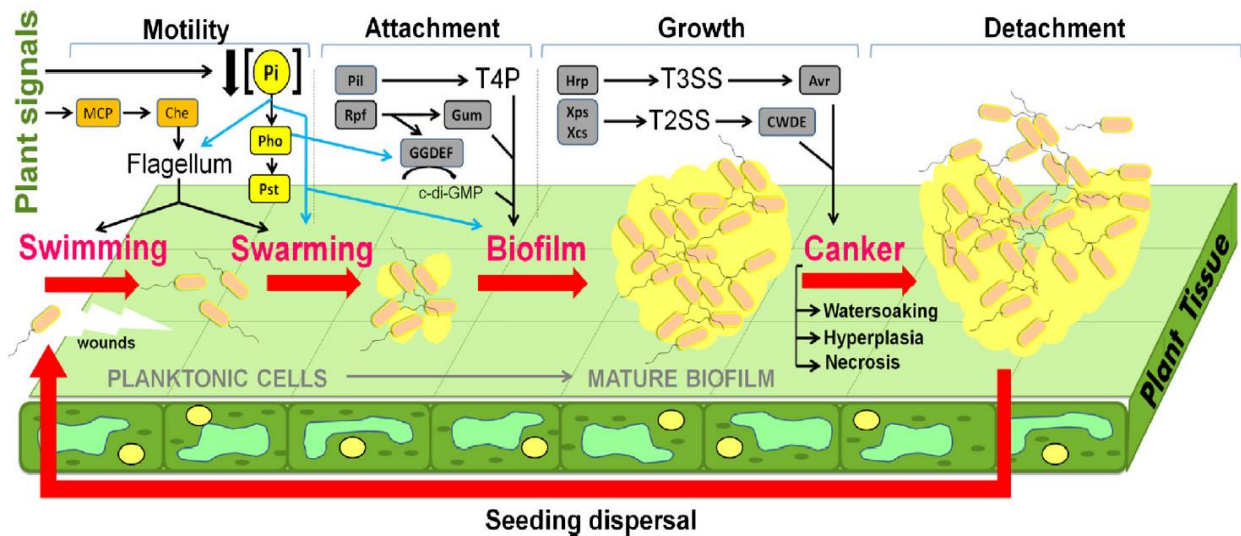
Figure 5. Colonization and biofilm formation of *X. citri* **during pathogen-host interaction.** Steps involved in colonization and biofilm are (1) plant signals, (2) swimming, (3) swarming, (4) biofilm formation and (4) canker generation. Flagellar movements could trigger swimming or swarming motility with the help of chemotactic proteins, favoring a possible shift to the interior of the plant. The low phosphate (Pi) concentration would signal, among other biological events; biofilm formation, growth, colonization of host tissues and canker induction. Blue lines highlight the possible relationships between phosphate internalization and storage with processes that contribute with bacterial virulence (Modified image, Source: MARCIO *et al.*, 2014).

1.2 Signal Transduction

Signal transduction refers to all biochemical processes which cells sense extracellular signals from their environment such as pH, nutrient levels, osmotic pressure, redox state, quorum-sensing proteins, antibiotics, etc., and evoke internal responses, which often leads to changes in gene expression and phenotypic alterations (TORRES AND FORMAN, 2006; KIEL *et al.*, 2010; TIWARI *et al.*, 2017). In the case of bacterial pathogens, they need to adapt to different microenvironments during interactions with the

23

host. In this respect, bacteria must sense and respond to extracellular signals to trigger the regulation of transport, metabolism, chemotaxis and osmoregulation (CRÉPIN *et al.*, 2011). Signaling processes include transient changes in the activities of enzymes and ion channels, altering steady-state concentrations of metabolites or ions, and posttranslational modifications (PTMs) such as protein phosphorylation, acetylation and glycosylation (TORRES AND FORMAN, 2006).

1.2.1 Phosphorus

The transport of phosphorus sources is essential for the growth of all living organisms. Phosphorus is an essential component of bacterial nutrition and is one of the major constituents of the cell, making up ~1.5–2.1% of the cell dry weight (SANTOS-BENEIT *et al.*, 2008). Phosphorous plays an important role in cell metabolism regulating proteins by the phosphorylation of specific amino acids like histidine, aspartate, serine, threonine, and tyrosine. Moreover, phosphorous make up the structural backbone of DNA and RNA, through their phosphodiester bonds, contributing to information storage and processing. Also, phosphorus is a biological buffer that contributes to pH homeostasis and assist in the coordination with metal ions, such as Mg^{2+} (SANTOS-BENEIT *et al.*, 2008; GARDNER *et al.*, 2019).

The preferred source of phosphorus in bacteria is the orthophosphate anion (PO_4^{3-}), most known as inorganic phosphate (Pi) (SANTOS-BENEIT *et al.*, 2015). Bacteria have mechanisms for sensing phosphate levels and altering the expression of numerous genes, many of which are involved in the acquisition and scavenging of phosphate more efficiently to contribute to maximal cell fitness (GARDNER *et al.*, 2019). The complete signal transduction systems that sense and transmits the inorganic phosphate (Pi) scarcity signal to the response regulator has not been yet completely elucidated and seems to vary from one organism to another (SANTOS-BENEIT *et al.*, 2015).

Essentially, bacteria have two different transport systems for Pi uptake defined by their affinity, low (Pit transporters) and high (Pst systems). Pit transporters or secondary transporters bring neutral metal-Pi complexes into the cell at the expense of a proton and they work when the environmental Pi levels are high. In some bacteria, these proteins

also can export the excess of intracellular Pi to maintain the cell homeostasis. PitA and PitB are the Pit transporters from *E. coli*. On the other hand, the high affinityPst system (PstSCAB) transporters are the principal importers under limiting Pi condition, and belong to the Phosphate regulon (Pho regulon), a global regulatory mechanism involved in bacterial Pi management (SANTOS-BENEIT *et al.*, 2008; MCCLEARY W. 2017; MICHIGAMI *et al.*, 2018).

1.2.2 Two-Component Systems and the Pho regulon

In bacteria, the major strategy of signal transduction is the two-component signal transduction system (TCS). A typical TCS is composed of two different types of proteins, a Histidine Kinase (HK) and a Response Regulator (RR) (LIU *et al.*, 2019). The receptor sensor is the HK which responds to external signals by activating the kinase domain, which autophosphorylates from an ATP at a conserved histidine residue. The phosphorylated sensor kinase interacts with the receiver domain of the response regulator (RR), which catalyzes the phosphoryl transfer to a conserved aspartate residue (MITROPHANOV AND GROISMAN, 2008). Phosphorylation of the RR activates its function of transcriptional regulator (MITROPHANOV AND GROISMAN, 2008; TIWARI *et al.*, 2017, LIU *et al.*, 2019) **(Figure 6A)**. TCSs participate in most aspects of bacterial physiology, including motility, sporulation, competence, nutrient uptake, stress response, central metabolism, and virulence (MONEDERO *et al.*, 2017). Several studies in different bacteria have shown that TCSs are essential for the expression of virulence factors, for bacterial viability and growth, for controlling biofilm formation and expression of antibiotic resistance determinants (TIWARI *et al.*, 2017; LIU *et al.*, 2019).

The PhoB/PhoR system constitutes the TCS that sense the phosphate limiting or starvation conditions in bacteria. In *E. coli* PhoB/PhoR is activated during extracellular phosphate concentration below 4 µM and induces the transcription of genes belonging to the Pho regulon (CRÉPIN *et al.*, 2011; VUPPADA *et al.*, 2018). PhoR is the sensor histidine kinase and PhoB is the response regulator. PhoR has a N-terminal domain anchored into the inner membrane, a PAS domain (Per, ARNT and Sim) that function to percept the signal, a DHp domain for **D**imerization and **H**istidine phosphorylation, and a

C-terminal CA domain that have the Catalytic and ATP-binding functions (VUPPADA *et al.*, 2018).

Although it remains unclear how PhoR perceives external Pi concentration, under Pi-limiting conditions, it autophosphorylates and transfers the phosphoryl group onto an aspartate residue of PhoB. In contrast, in Pi-rich conditions, PhoR dephosphorylates phospho-PhoB through its phosphatase activity. PhoU, the negative regulator of the Pst system participates in the PhoR activity (CRÉPIN *et al.*, 2011; VUPPADA *et al.*, 2018; GARDNER *et al.*, 2019). Upon phosphorylation, PhoB dimerizes in the receptor (REC) domain and acts as an active transcription factor that binds in the PHO box region located in the Pho regulon operons and genes, positively regulating them **(Figure 6B)** (VUPPADA *et al.*, 2018; RIGHETTI *et al.*, 2021).
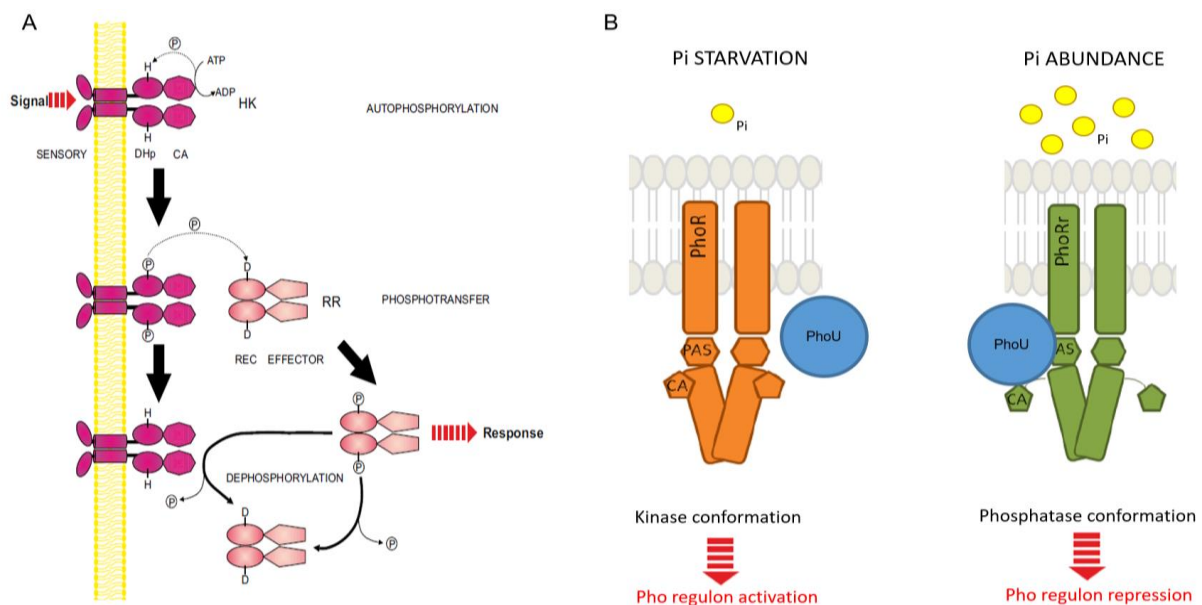
Figure 6. Two-component system (TCS) conformation. A. In general the TCS consists of two proteins: a sensory protein called Histidine kinases (HK) and a response regulator protein (RR). The signal transduction involves three steps: first, autophosphorylation of the HK using an ATP upon detection of a specific signal, specifically the DHp domain is phosphorylated at a conserved Histidine (H) residue; second, the phosphate transference from the phosphorylated histidine to the Aspartate (D) residue of the REC domain in RR protein; and third, conformational change of RR by phosphorylation or dephosphorylation regulates the RR activities as phosphatase or autokinase, respectively. In addition, RR phosphorylated binds to DNA and acts as transcriptional factor. **B.** The response mechanism of PhoBR TCS to different environmental conditions. The structural changes in PhoR induced by Pi starvation (orange) and Pi excess (green). Note that PhoU (blue) plays a central role in the kinase activity. When bound to PhoR through both the PAS and CA domains leads to PhoR phosphatase conformation, and when PhoU is not bound, PhoR remains with the kinase conformation (Image source: MONEDERO *et al.*, 2017; RIGHETTI *et al.*, 2021).

PHO box is a specific consensus sequence that varies from one species to another and leads to different affinities of PhoB protein (SANTOS-BENEIT *et al.*, 2015; MICHIGAMI *et al.*, 2018). For instance, the most conserved motif of PHO box in *E. coli* is CTGTCAT and in *Streptomyces coelicolor* is GTTCACC. In *E. coli*, the interaction PhoB/PHO box occurs in at least 22 promoters and influences the level expression of more than 400 proteins (GARDNER *et al.*, 2019), some of these proteins with pleiotropic effect (CRÉPIN *et al.,* 2011).

There are additional features associated to the PHO boxes and involved in the control of the PHO regulon. One of these features is the sigma (σ) factor or σ subunit of the RNA polymerase (SANTOS-BENEIT *et al.*, 2015). In *E. coli*, the primary σ factor is σ^{70}, there is crystallographic evidence that PhoB directly contacts the C-terminal region of σ^{70} to accomplish RNA polymerase recruitment and trigger the transcription initiation (BLANCO *et al.*, 2011; SANTOS-BENEIT *et al.*, 2015).

1.2.3 Pst system and ABC transporters

One of the most important systems regulated by PhoB protein is the ATP-Binding Cassette Pst system (*Phosphate specific transport system*), a Pi transporter of high affinity (ABC transporter) (CRÉPIN *et al.*, 2011; SCHRAMKE *et al.* 2017).

ABC transporter superfamily translocate diverse range of substrates and can be importer or exporters (THOMAS *et al.*, 2020). Structurally, these complexes in the inner membrane consist of two nucleotide-binding domains (NBDs or ATPases) that gives energy to the transport through the ATP breakdown and two transmembrane domains (TMDs) that form a pore for the ligand passage and, in case of importers, a periplasmic substrate-binding protein (SBPs) that is responsible for the affinity and specificity of the transporter (BEEK *et al.*, 2014; WILKENS *et al.*, 2015; BERNTSSON *et al.*, 2010).

The NBD is the most conserved component in the ABC transporters and has the ABC signature, motif LSGGQ. The transport occurs after ATP binding and hydrolysis, which drives structural changes in the TMD, resulting in alternating access from inside and outside of the cell for unidirectional transport across the lipid bilayer (BEEK *et al.*, 2014; WILKENS *et al.*, 2015).

ABC transporters were firstly divided in four types depending on their architecture and direction of the substrate. Three types of ABC transporter were associated with import functions: Type I and Type II importers and the Energy Coupling Factor (ECF) transporters (also named Type III importers). The Type IV was characterized as exporters (BEEK *et al.*, 2014; LOCHER K. 2016). In 2020, Thomas and collaborators (THOMAS *et al.*, 2020) proposed a new classification of ABC transporters based on TMD sequence homology and architecture. Seven distinct folds or types were suggested: Type I to III describe importer transporters, Types IV and V are exporters with some notable exceptions, Type VI are extractors, and Type VII are components of mechanotransducing tripartite efflux pumps, or operate outside of tripartite assemblies as lipoprotein extractors, or still as cell division regulators (THOMAS and TAMPÉ, 2020). The Pst system is encoded by the *pst* operon (*pstSCAB-phoU*) (JACOBSEN *et al.,* 2008; TETSCH *et al.,* 2009; HADDAD A. 2010; MCCLEARY W. 2017) and composed by a homodimer of PstB (NBDs), a heterodimer of PstA and PstC proteins (TMDs) and the periplasmic SBP PstS. All together, these proteins define a Type II importer (LOCHER K. 2016; THOMAS *et al.*, 2020). The product of *phoU* gene is a negative regulator of the transport that also mediates the interaction between the Pst and the PhoR/PhoB systems (COX *et al.,* 1989; KURODA *et al.,* 2000; DÍAZ *et al.,* 2005; FURTWÄNGLER *et al.,* 2010; PEGOS *et al.,* 2014). In the same operator, Structurally, Pst system can adopt at least two conformational states during the transport process, inward with NBDs apart or outward-facing with both NBDs close and bound to ATP (ULUŞEKER *et al.*, 2019; MARTÍN *et al.*, 2021). In excess of Pi, PhoU is bound to PstB negatively controlling the transport and PhoR autokinase activity. On the other hand, in low levels of the anion, the phosphate bound PstS anchors in the periplasmic interface of PstA and PstC, triggering structural changes in the transmembrane proteins that stimulate the ATP hydrolysis and the consequent phosphate transport (MICHIGAMI *et al.*, 2018; MARTÍN *et al.*, 2021. PhoU then favors PhoR autophosphorylation. When TCS PhoB/PhoR is active, the Pst transporter is up regulated but otherwise, in the presence of high levels of Pi, TCS PhoB/PhoR is inactive and Pst transporter is down-regulated (JACOBSEN *et al.,* 2008; TETSCH *et al.,* 2009; HADDAD A., 2010). The sensitive tunning of phosphate regulation

is necessary to maintain the anion intracellular levels within an optimal range and to avoid the Pi poisoning (JACOBSEN *et al.,* 2008; MCCLEARY W. 2017) **(Figure 7).**
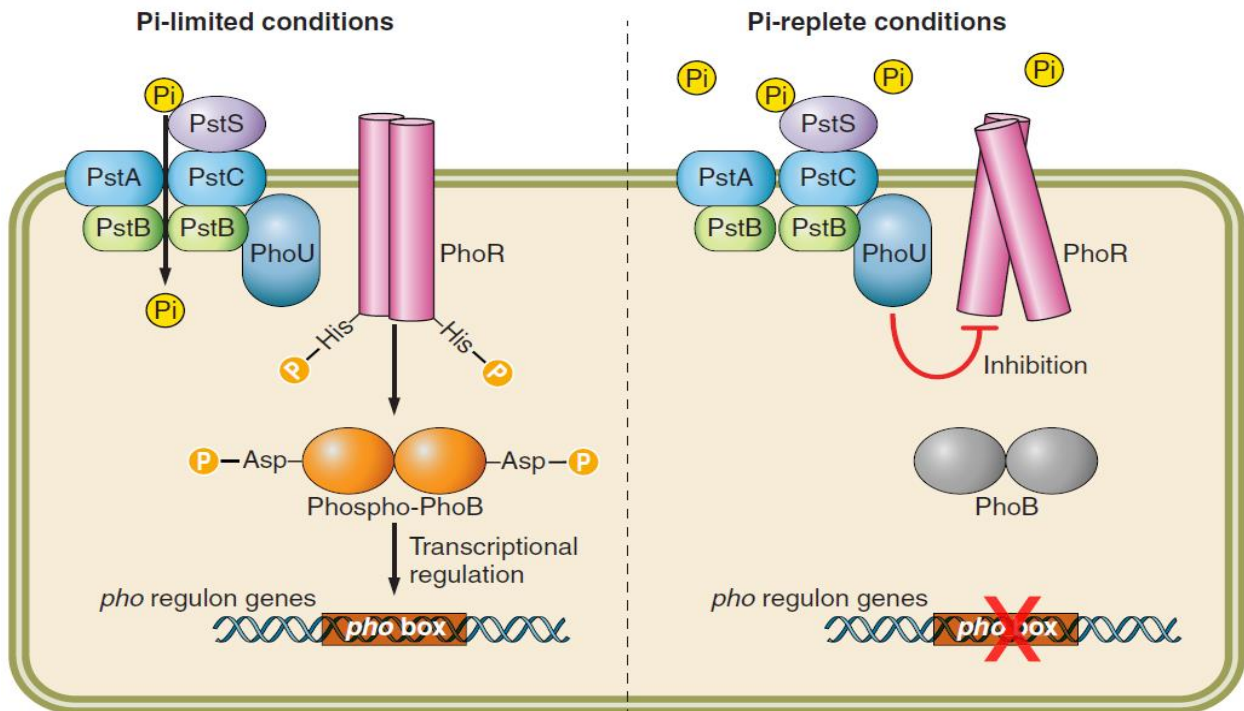
Figure 7. The TCS PhoB/PhoR and the Pst system activities for Pi transport. The Pst system consists of the periplasmic-binding domain, PstS, the transmembrane domains, PstA and PstC, and the intracellular nucleotide-binding domain, PstB. According to the current model, PhoR assesses Pi availability by monitoring the activity of Pst transporter, and relays the signal from PstB via PhoU to PhoR. When external Pi is in excess, PhoU stabilises PhoR and prevents PhoR from autophosphorylation. Consequently, the transcription factor PhoB does not become phosphorylated by PhoR and is not able to bind the DNA in the PHO boxes activating the transcription of Pho genes. When external Pi is limited, PhoU favors the PhoR autokinase conformation allowing the phosphorylation of PhoB, its dimerization and recruitment of σ factors from the RNA polymerase to finally bind to the PHO boxes in the DNA. The up-regulation of *pst* genes and consequent Pi transport occurs (Image source: MICHIGAMI *et al.*, 2018; ULUŞEKER *et al.*, 2019).

1.3 Pst system and phosphate uptake in *Xanthomonas citri* subs. *citri*

X. citri, similarly to what is observed in *E. coli*, has *pstSCAB-phoU* and *phoRB* operons encoding the Pst system and PhoU and the TCS PhoR/PhoB, respectively. In addition, *X. citri*, has the operon *phoXoprO* (XAC1578/XAC1579), which encodesa PstS paralog that binds Pi and a polyphosphate-selective porin O, respectively.

Phylogenetic analysis comparing PhoX and PstS with orthologues in different species of Proteobacteria, concluded that both proteins are originated from a gene

duplication occurred in an ancestral orthologue from *Desulfatibacillum alkenivorans* (PEGOS *et al.*, 2015; 2017). The presence of two phosphate-binding proteins is also observed in some species of *Xanthomonas* genus: *Xanthomonas axonopodis* pv. citrumelo F1, *Xanthomonas campestris* pv. vesicatoria and *Xanthomonas campestris* pv. campestris ATCC 33913.

Other Pho regulon genes were identified in the *X. citri* genome, encoding three putative alkaline phosphatases that increase the availability of Pi from organophosphates, *pho*A (XAC2759), *pho*X (XAC1792) and *phoD* (XAC4166). Similarly, genomic analysis revealed *phoH* (XAC2462), *ppx* (XAC1039) and *ppk* (XAC1039) genes that encode a predicted phosphate starvation-inducible protein, a exopolyphosphatase and a polyphosphate kinase, respectively **(Figure 8)**.
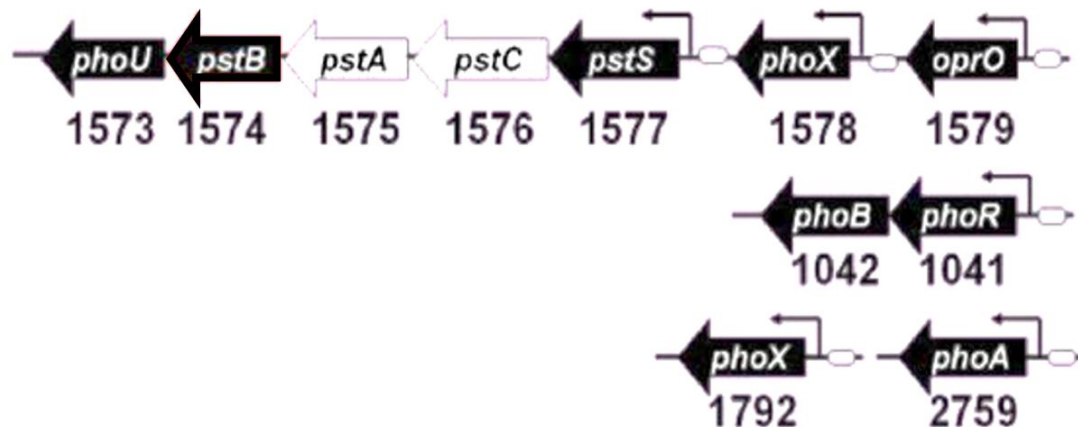
Figure 8. *X. citri* operons and genes involved in the phosphate transport by Pst system. The *pstSCAB-phoU* and *phoB/pho*R operons and the genes *pho*X (XAC1578), *opr*O (XAC1579), *pho*X (XAC1792) and *pho*A (Image source: MARCIO *et al.,* 2014).

All the mentioned proteins are expressed in *Citrus sinensis* and *Citrus limonia* leaves infected with *X. citri*, highlighting the relevance of the Pho regulon genes during infection and growth **(Figure 9)** (MARCIO *et al.,* 2014).
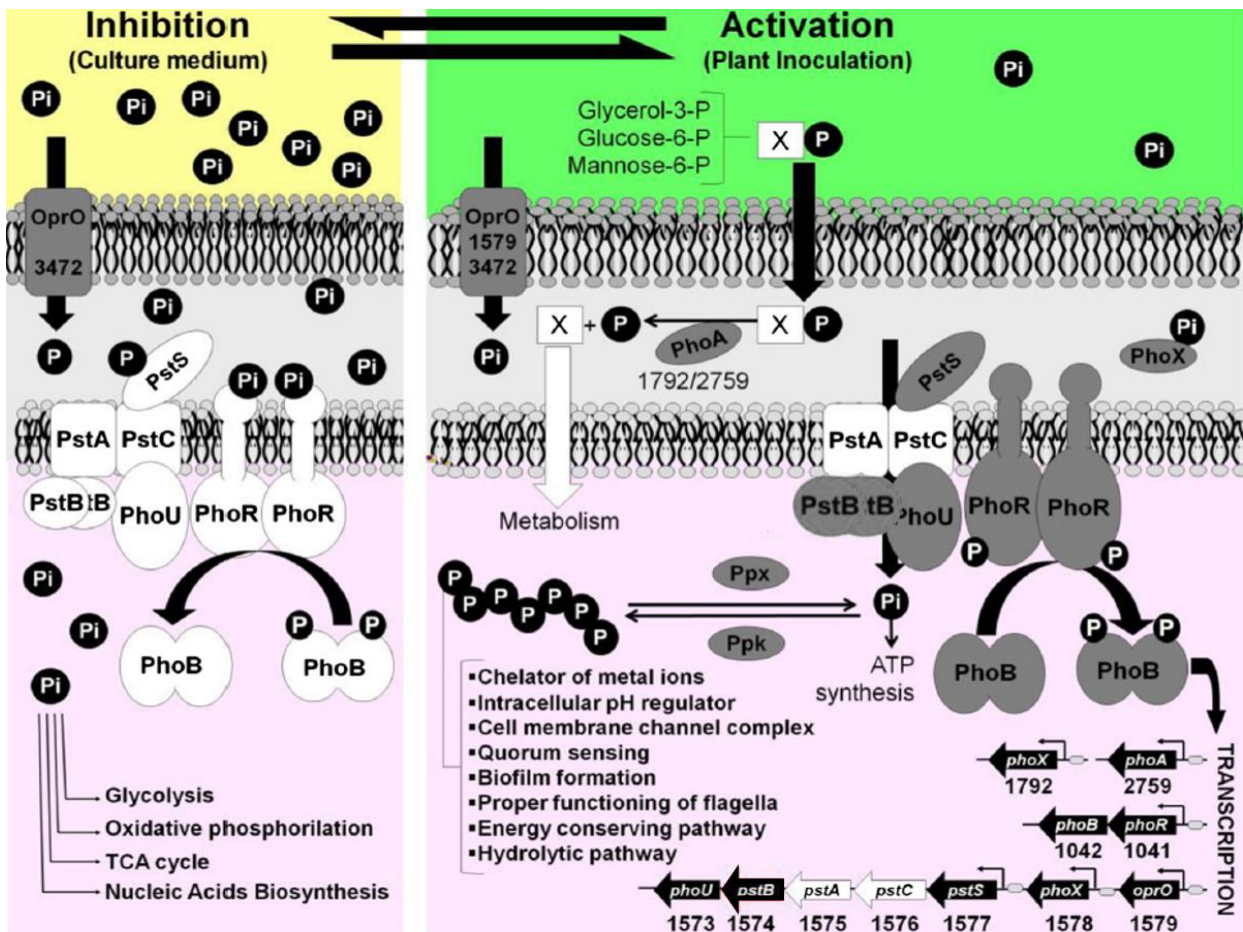
Figure 9. Regulation model of the genes involved in the phosphate uptake and transport by Pst system in *X. citri*. In culture medium with a high concentration of inorganic phosphate (Pi), *X. citri* suppresses or inhibits the TCS PhoB/PhoR negatively regulating the synthesis of proteins that composed the Pst system and other proteins related to the *pho* regulon. In opposition, low levels of phosphate during *X. citri* infection in plants activates the TCS PhoB/PhoR triggering the expression of proteins of the Pst system (PstSCAB-phoU), porines and proteins associated to phosphate metabolism and storage (alkaline phosphatases- PhoA, exopolyphosphatase- Ppx and polyphosphate kinase- Ppk) (Image source: MARCIO *et al.,* 2014).

1.3.4 *PstS and PhoX as Substrate Binding Proteins (SBPs)*

The substrate-binding domains (SBDs) are free in the periplasm of Gram-negative bacteria and lipid-anchored or fused to the TM in the case of Gram-positive bacteria and Archaea. SBPs are associated with a wide variety of complexes and play an important role in signal transduction, solute uptake and deliver of molecules into the cell for translocation by the TMDs (BERNTSSON *et al.*, 2010; SCHEEPERS *et al.*, 2016). Despite the low sequence similarity among SBPs of different ABC importers, they share a common

architecture comprising two structurally conserved rigid lobes connected by a flexible hinge region. Biophysical and structural studies indicate that ligand binding site is located at the interface between two lobes, which promotes the switch between two conformations, the open (unbound) and closed (bound) (DE BOER *et al.*, 2019). Firstly, SBPs were classified in two types, I and II, based on the secondary structure elements of the hinge regions (FUKAMI-KOBAYASHI *et al.*, 1999). This classification was reviewed in 2010 by Berntsson and collaborators that suggested six different clusters, from A to F, based on the available three-dimensional structures (BERNTSSON *et al.*, 2010).

X. citri PhoX (XAC1578) and PstS (XAC1577) are located in the periplasm as monomers that belong to class II (two connecting strands in the hinge region) and cluster D (two short hinges) **(Figure 10)** (BERNTSSON *et al.*, 2010).
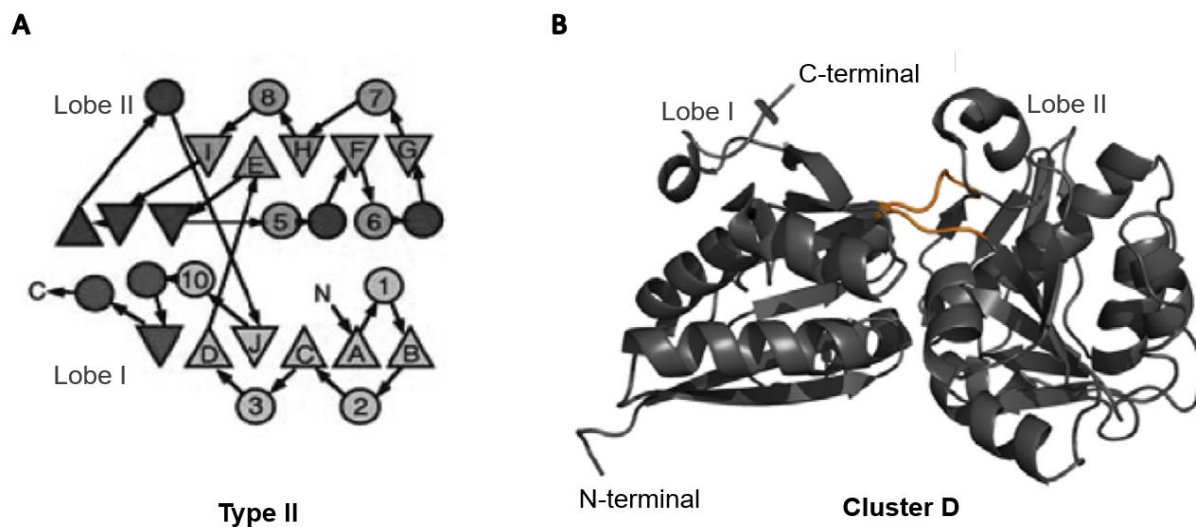
Figure 10. Characteristics of the SBPs belonging to Type II and Cluster D. A. The topology arrangement diagram of SBPs Type II consist of two lobes connected by two strands (the hinge region). In the diagram helices are shown as circles and strands are shown as triangles. **B.** SBPs from Cluster D contain two relative short hinges in orange color between the two lobes (Image source: FUKAMI-KOBAYASHI *et al.*, 1999; BERNTSSON *et al.*, 2010).

1.3.5 PstS and PhoX as Phosphate Binding Proteins (PBPs)

Phosphate binding proteins (PBPs) are members of the SBPs that bind phosphate with high-affinity (dissociation constants reach the sub-micromolar range) and high-specificity against competing anions (YAO *et al.*, 1996; GONZALEZ *et al.*, 2014; ELIAS *et al.*, 2012). Proteomic analysis of *X. citri* revealed that PstS and PhoX are highly induced

under Pi limited conditions, with fold change values of 43.5 and 39.1, respectively (PEGOS *et al.*, 2014). Both proteins are encoded from genes located in different operons that present a PHO-box in the promoter region. Indeed, in *X. citri* mutant strain deleted for *pst*S gene (*Xac::pstS*),), PhoX was capable to complement the Pi uptake, which includes: phosphate bind by PBPs, the subsequent phosphate release within PstA and PstC permeases and the phosphate transport through permeases into the cell. However, PstS seemed to play some additional regulatory function, since its absence leaded the constitutive expression of the *pstSCAB* operon and significant differences in the profile of expressed proteins and metabolites in the cell **(Figure 11)** (PEGOS *et al.*, 2015).
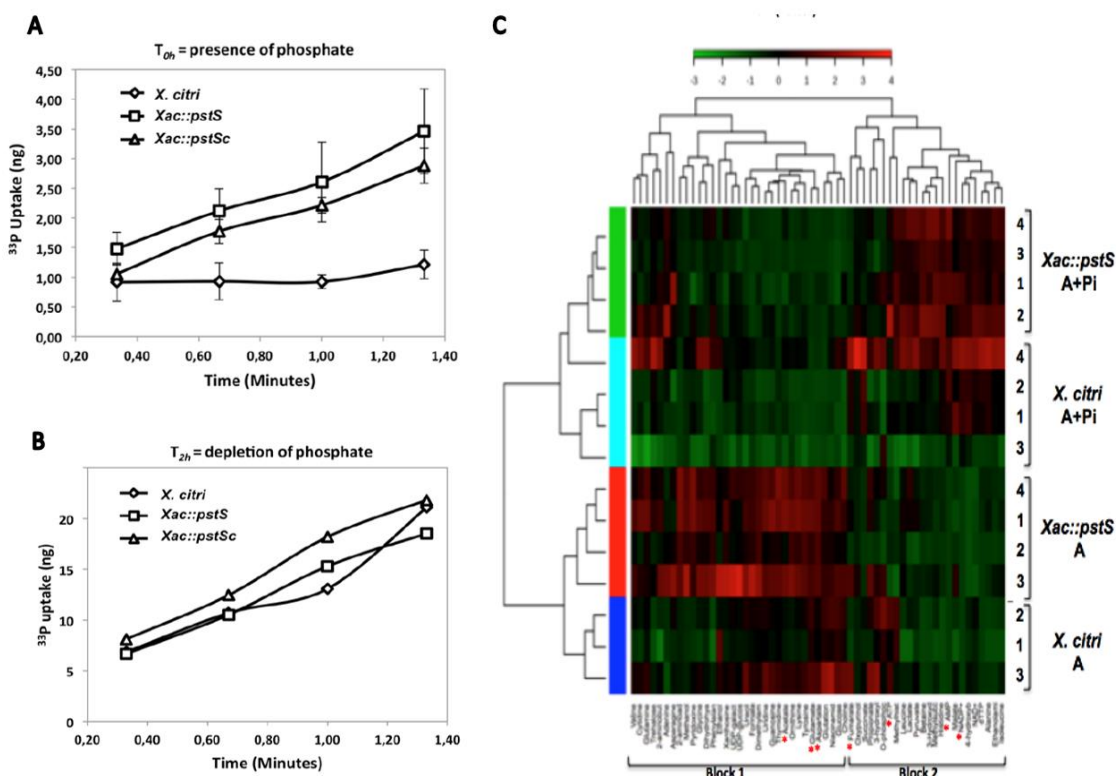
Figure 11. Effects of *pst*S deletion in the phosphate uptake and transport, and also in the transcriptional regulation of *X. citri.* A. *X. citri, Xac::pstS* and *Xac::pstSc* (complementary) strains were growth on T-salt medium supplemented with 4 mM of KH_2PO_4 and 0.2% of glucose, until OD_{600nm} = 1.0 . Cultures were washed and resuspended in Pi-free medium plus radioactive phosphate label (radioactive ^{33}P). Plot A used data immediately collected after washing and inoculation in presence of phosphate (T_{0h}). **B.** After procedure described in A, Plot B used data collected after two hours of incubation at 28°C (T_{2h}) in depletion of phosphate. **C.** Heat map of the metabolomes of *Xac* wild type and *Xac::pstS* strain. The plot is shown in a cladogram that separates A (*Xac* cells grown in minimal medium) and A+Pi (*Xac* cells grown in minimal medium supplemented with 10 mM of Na_2HPO4) results. Red and green colors correspond to high and low levels of identified metabolites, respectively. (Image source: Pegos *et al.*, 2015).

Structurally, all phosphate-binding proteins (PBPs) with solved structure possess the same folding consisting of a venus fly-trap topology (RAVI *et al.*, 2007) **(Figure 12A)**, but so far, there are no structures of these proteins in the open form (without ligand and with Domains I and II separated). The bound structures revealed that upon ligand binding, the interface of both domains lies a phosphate-binding cleft, which comprises 8 residues, forming a set of hydrogen bonds with the phosphate molecule in several PBPs **(Figure 12B)** (LEDVINA *et al.*, 1998; ELIAS *et al.*, 2012; GONZALEZ *et al.*, 2014). In addition, the bound anion is also immobilized by electrostatic interactions (including van der Waals forces and dipolar interactions). Most of the PBPs have a key aspartate residue in Domain I, responsible for the low barrier hydrogen bond (LBHB) with the bound phosphate anion and the extremely high selectivity over various competing anions (LEDVINA *et al.*, 1998; GONZALEZ *et al.*, 2014). Moreover, Domain II has an arginine residue that seem to favor a docking site with an aspartate that is potentially susceptible to the changes observed with increasing ionic strength **(Figure 12C)**. Access to and from the cleft is accompanied by a hinge bending between the two domains. PBP binds both monobasic and dibasic phosphates in a broad pH range (from 4.3 to 9) (LEDVINA *et al.*, 1998). The complex PBP-Pi is subsequently transported across the membrane by Pst ABC transporter against the concentration gradient (YAO *et al.*, 1996; ELIAS *et al.*, 2012).

The PBPs, PstS and PhoX from *X. citri* have an amino acid sequence identity of 70%. In a previous work (Pegos *et al.,* 2017), the structure of PhoX was solved in a closed conformation with a phosphate engulfed in the binding site pocket between two domains, confirming its function as a phosphate-binding protein **(Figure 13A)**. As reported by the authors, the phosphate pocket in PhoX is coordinated through 12 hydrogen bonds performed by seven residues: Ser34, Ser62, Asp80, Arg159, Gly164, Thr165 and Asn201 **(Figure 13B)**, all conserved in PstS: Ser58, Ser86, Asp104, Arg183, Gly188, Thr189 and Asn225.

Another important characteristic of PhoX and PstS is the conservation of residues in the two regions of interaction with PstC and PstA permeases, and denominated RI and RII for Domains I and II, respectively. Although these regions are similar in both proteins, residues belonging to RI are more conserved than those present in RII **(Figure 14A)** and

differences in the electrostatic potential are evidenced, suggesting that PhoX and PstS might have different affinities for the permeases **(Figure 14B)**.
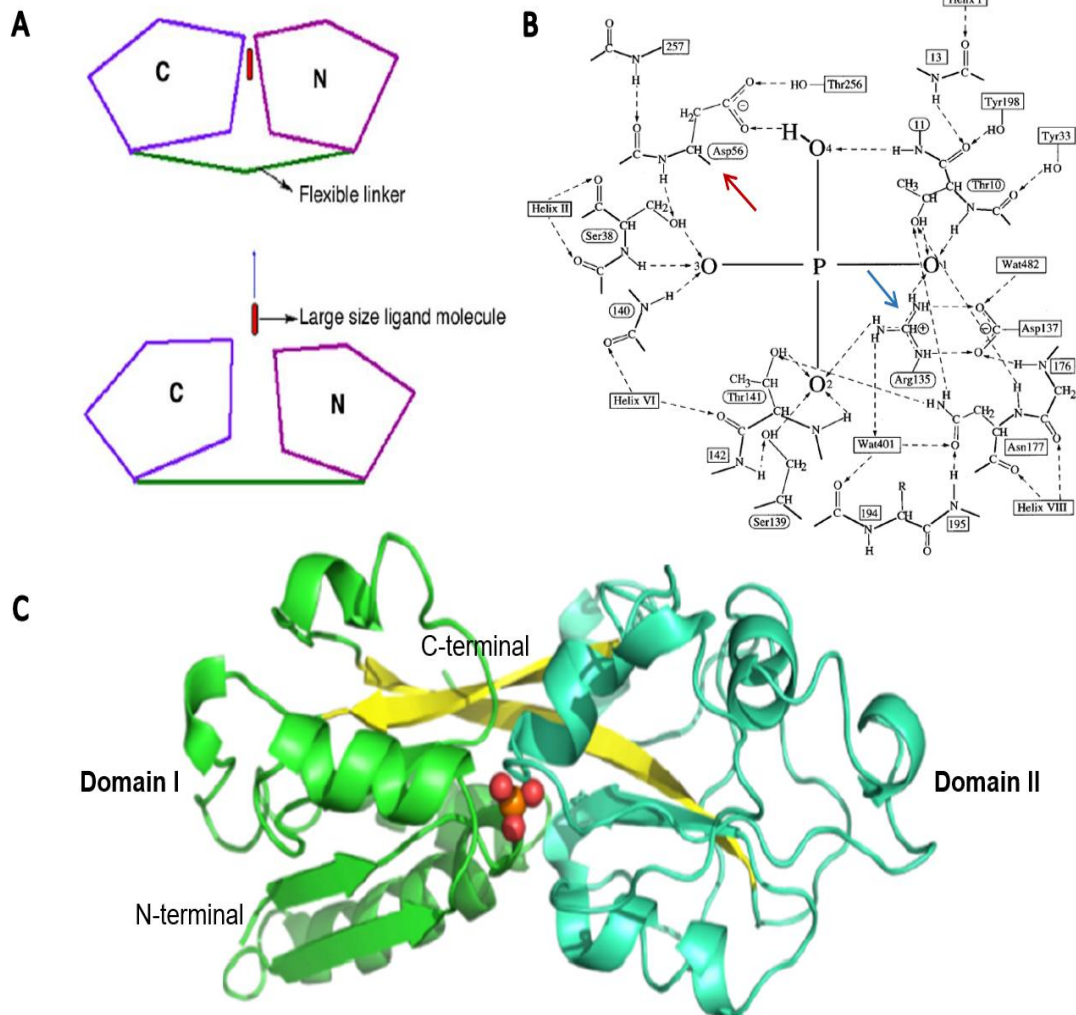
Figure 12. Structural characteristics of PBPs. A. PBPs have a Venus fly-trap topology which involves a mechanism for capture/release of ligands. The N and C-domains are held by a flexible linker which acts as a hinge. Bending of the flexible linker allows opening of the two domains so the trapped ligand can escape. **B.** Schematic diagram of the 12 hydrogen bonds (broken lines) and 1 ionic interaction typically present when phosphate binds a wild-type PBP. The red arrow points to the residue Asp56, which is negatively charged, and confers selectivity to the ligand while the blue arrow points to the residue Arg135 that could form an ionic bind with Asp137. **C.** Domains of PBPs (Domain I and Domain II) are colored in green and cyan, beta-sheets across the two domains in yellow and the phosphate between both domains is in red/orange spheres (Image source: LEDVINA *et al.*, 1998; GONZALEZ *et al.*, 2014; RAVI *et al.*, 2007).
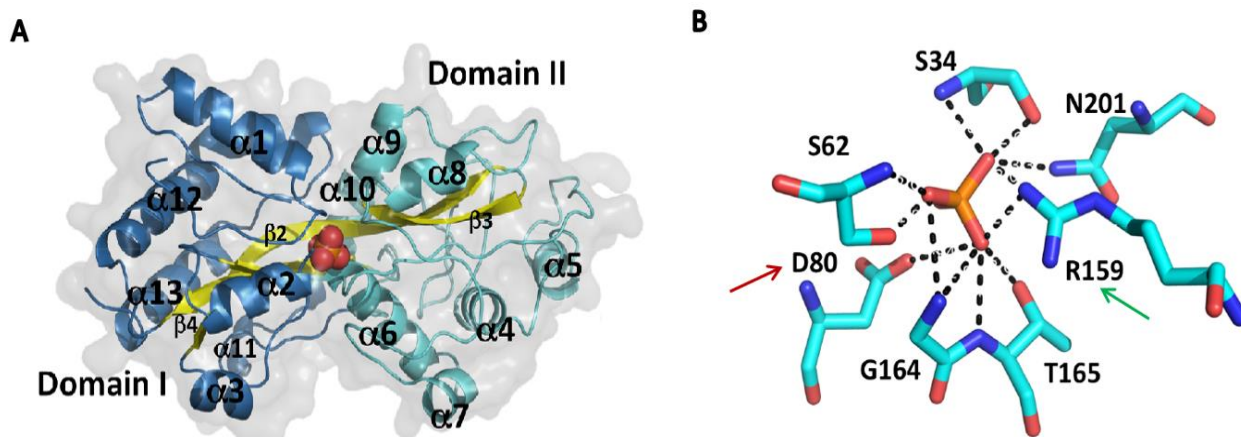
Figure 13. Structural characteristics of *X. citri* PhoX. A. Cartoon representation of the three-dimensional structure of PhoX evidencing the presence of two domains, Domain I and Domain II, in blue and cyan color, respectively. The phosphate ion is shown as red spheres and the sheets connecting the two domains is colored in yellow (Image source: PEGOS *et al.*, 2017). **B.** Binding pocket of PhoX is composed by seven residues (S34, S62, D80, R159, G164, T165 and N201) and 12 hydrogen bonds (broken lines). It is predicted that residue D80 confers specificity to the ligand and residue R159 could form ionic bonds with other aspartate residue stabilizing the pocket (Image source: PEGOS *et al.*, 2017).
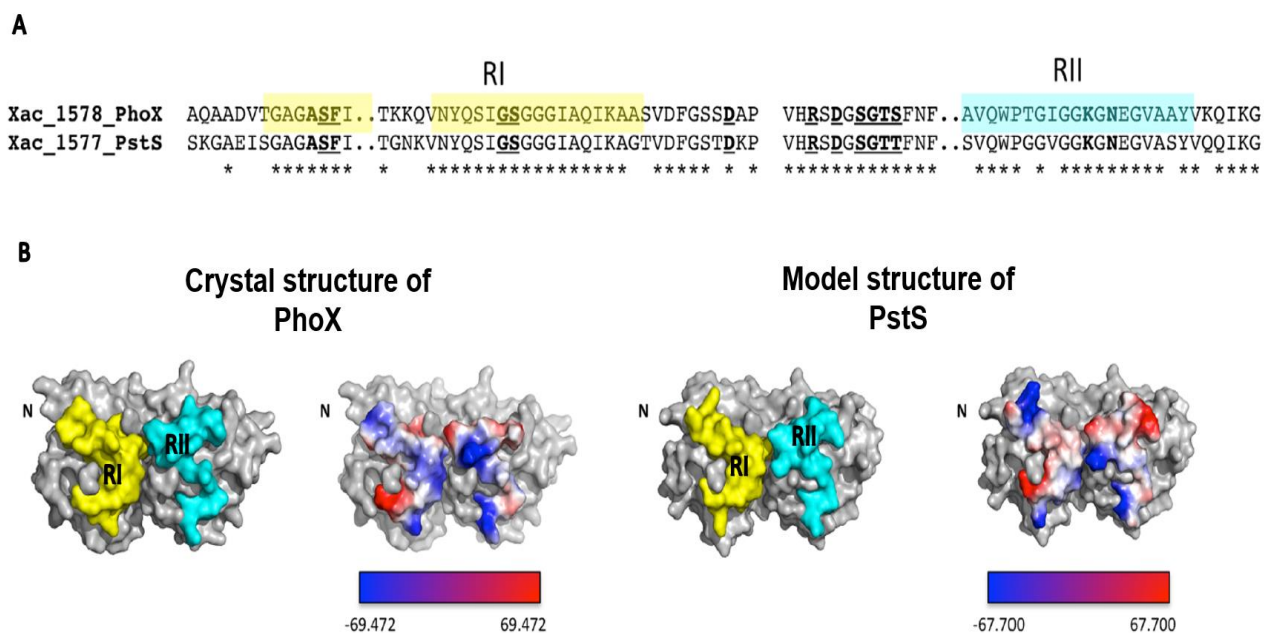
Figure 14. Comparison of *X. citri* PhoX and PstS in the region of interaction with permeases PstAC. A. Amino acid sequence alignment of RI and RII from PhoX and PstS evidences a high level of conservation (asterisks). Residues in underlined bold are involved with the phosphate binding. **B.** Surface representation of the three-dimensional structure of PhoX and PstS locating the regions RI (yellow) and RII (cyan) which interact with the permeases PstCA are represented on the left. On the right are shown the electrostatic potential of RI and RII of both proteins, colored according to the charges calculated in Pymol. Negative: red, positive: blue and neutral: white (Image source: PEGOS *et al.*, 2017).

36

So far, some questions about the functional differences between PstS and PhoX remain to be answered such as: (i) what about the phosphate affinity of PstS and PhoX?; (ii) there are differences between PstS and PhoX regarding the mechanism of phosphate uptake?; (iii) could PstS and PhoX induce different structural conformations in the transporter? (iv) could PstS and PhoX be capable of interacting with other proteins in *X. citri* as it happens in *Pseudomonas aeruginosa*? The results obtained until now so far suggests that although PhoX be capable to complement PstS transport function, PstS has additional functions in the cell. To answer these questions and to establish the functional differences between PstS and PhoX more studies are necessary.

1.4 *X. citri* chemotactic proteins and its possible relation with PBPs

Chemotactic proteins are responsible for bacteria movement toward or contrary a gradient concentration in presence of chemoattractants or chemorepellents. These proteins are mainly involved in bacteria adaptation to different environmental conditions (SOARES *et al.,* 2010). In plant pathogens, like *X. citri*, chemotaxis can be led by roots exudates or secretions from plant lesions (MARCIO *et al.*, 2014).

Methyl-accepting chemotaxis proteins (MCPs) are the principal sensory receptors of the bacterial chemotaxis system, these proteins receive a particular stimulus to generate a signal that leads a cellular response (STOCK and BAKER, 2009). There are five MCPs types which different ligand specificities: Tsr (taxis to serine and repellents), Tar (taxis to aspartate and repellents), Tap (taxis to dipeptides), Trg (taxis to ribose and galactose) and Aer (aero taxis), the latter does not have a periplasmic domain but have an N-terminal Per-Arnt-Sim (PAS) domain, which binds with a flavin adenine dinucleotide to sense redox changes (BI and LAI, 2015; PARKINSON *et al.*, 2015). For most bacteria, including *X. citri,* Tsr proteins are the abundant group within MCPs (MARCIO *et al.*, 2014; MOWERY *et al.*, 2015).

MCPs have a consensus structure comprising three domains: a sensory domain or receptor, a transmembrane domain and a signaling domain. Moreover, the signaling domain also harbors two conserved domains: a histidine kinase, adenyl cyclase, methyl-

accepting chemotaxis protein and phosphatase domain (HAMP) and a MCP domain (MARCIO *et al.*, 2014; SALAH *et al.*, 2017) **(Figure 15A)**. The only hypervariable region is the sensory domain, which is located in the periplasm and is responsible for the perception of extracellular signals that lead to the activation of the flagellar response upon activation (by phosphorylation) of the *che* genes (SALAH *et al.*, 2017). Chemotactic proteins of most bacteria are composed of two transmembrane domains, whereas only one transmembrane domain is observed in *X. citri* **(Figure 15B)** (MARCIO *et al.*, 2014).
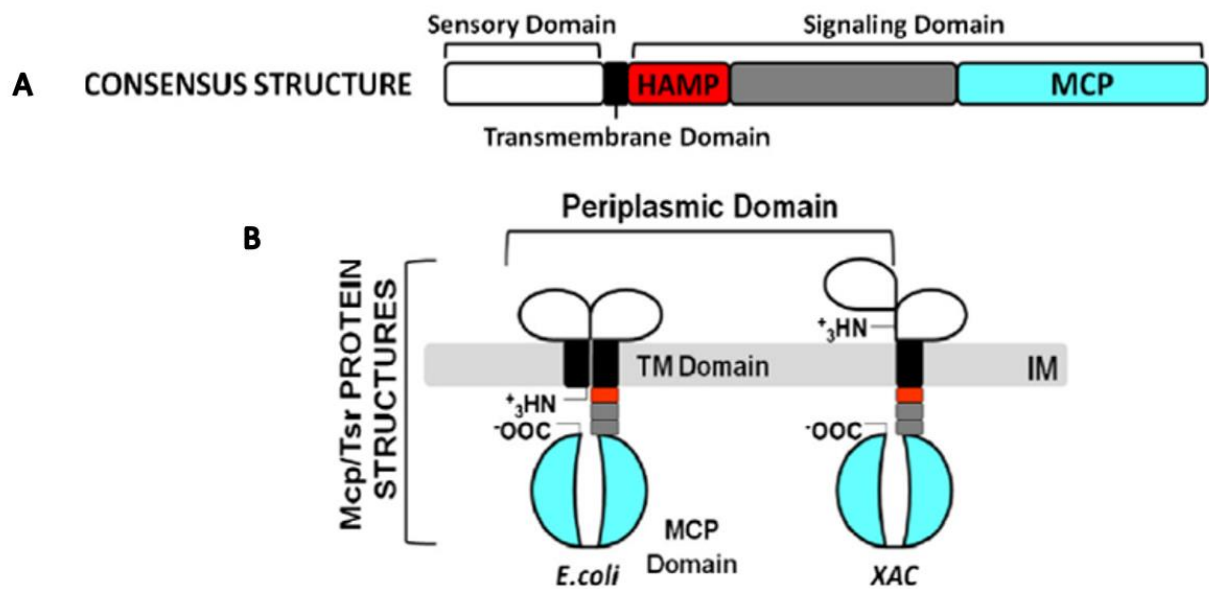
Figure 15. Consensus structure of chemotatic proteins in *E. coli* **and** *Xac***. a.** Consensus structure of chemotactic genes are composed by three domains: sensory, transmembran and signaling domains. **b.** Differences between the structures of chemotactic proteins (Mcp and Tsr) of *E. coli* and *Xac*. In *E. coli*, chemotactic proteins haves two transmembrane domains whereas *Xac* have only one transmembrane domain (Image source: Marcio *et al.*, 2014).

Sensory domains of chemotactic proteins recognize their ligands directly or by the interaction with periplasmic binding proteins (BI *et al.*, 2015; SALAH *et al.*, 2017). This can be exemplified in the study of RICO-JIMÉNEZ *et al.*, 2016 where the chemotactic proteins: CtpH (PA2561) and CtpL (PA4844) from *Pseudomonas aeruginosa* (*P. aeruginosa*) can interact with Pi or PstS and lead chemotaxis toward high and low Pi concentration, respectively. These chemotactic proteins use different mechanisms: either by direct recognition of Pi (as CtpH) or by binding to PstS attached to Pi (as CtpL). Both

proteins have a specific region called ligand binding domains (LBD) that recognize specific chemoeffectors (as Pi or PstS) leading to the activation of the sensory domain of chemotactic proteins. Typically, LBDs are composed by amino acids of the periplasmic domain of a protein (LACAL *et al.,* 2010) and the affinities for ligands or PstS can be different (RICO-JIMÉNEZ *et al.,* 2016).

The available data of *X. citri*, PstS and PhoX show that although they are highly conserved in structure and residues for phosphate-binding, they also have significant differences that point for additional roles in the bacterium, such as differences in the expression profiles in presence and absence of phosphate and completely altered metabolism in absence of *pstS*. More functional and structural studies of chemotactic proteins from *X. citri* are necessary to determine their implication in phosphate signaling and its possible interaction with chemotactic proteins.

2 OBJETIVES

Principal objective

To characterize the structural behavior of PstS and PhoX from *Xanthomonas citri* subsp. *citri* in the uptake of phosphate and to establish their relationship with chemotactic proteins.

Specific objectives

Chapter I

- Expression and purification of the PstS and PhoX proteins;
- Determination of the affinity of PstS and PhoX for phosphate;
- Evaluation of the PstS and PhoX behavior during molecular dynamics simulations in presence and absence of phosphate;

Chapter II

- Expression and purification of the chemotactic proteins Tsr1 and Tsr2;
- Construction of *X. citri* mutant strains deleted for *phoX* (*Xac::phoX*);
- Determine the relationship between PstS and PhoX with Tsr1 and Tsr2 of *X. citri*.

3 MATERIALS AND METHODS

3.1 Chapter I

3.1.1 Material, bacterial strain and plasmids

Pi refers to K_2HPO_4 (Sigma). The strain used was *Escherichia coli* BL21(DE3) and the plasmids used were pET28a$^+$-PhoX and pET28a$^+$-PstS.

3.1.2 Expression, purification and western blot of PhoX and PstS

PhoX and PstS proteins from *Xanthomonas citri* subsp. *citri* 306 (*X. citri*) were expressed and purified as described previously in Pegos and collaborators (2014; 2017), with some modifications. Briefly, plasmids pET28a$^+$-PhoX and pET28a$^+$-PstS **(Appendix 1A, 1B)**, containing the nucleotide sequences corresponding to the mature proteins PhoX and PstS in fusion with a N-terminal His-tagged were acquired from the mentioned studies. These plasmids were transformed in *E. coli* BL21(DE3) cells and cultured in 2 liters of LB medium (10 g.l^{-1} of peptone, 10 g.l^{-1} of yeast extract and 5 g.l^{-1} of sodium chloride) supplemented with 50 µg.ml^{-1} of kanamycin at 37°C. Expression of recombinant proteins were induced up to OD600 of 0.6 by adding 0.1 mM of isopropyl-β-D-1-thiogalactopyranoside (IPTG) and keeping growing for 4h at 30°C. Cells were harvested by centrifugation and resuspended in lysis buffer (50 mM Phosphate buffer pH 7.2, 100 mM NaCl, 20 mM imidazole and 5% of glycerol) and then lysed on ice by sonication. After centrifugation, both proteins were first purified by immobilized nickel-affinity chromatography (IMAC) using a HisTrap Chelating HP column (GE Healthcare Life Sciences) previously equilibrated with Elution Buffer A (50 mM Phosphate buffer pH 7.2, 100 mM NaCl and 20 mM imidazole). Bound proteins were eluted in a range between 68- and 180- mM of imidazole in Buffer A. Furthermore, the proteins were further purified by size-exclusion chromatography (SEC) in Hi Load 16/60 200 Superdex column (GE Healthcare Life Science) with 10 mM Tris-HCl pH 7.5 elution buffer and 10 mM NaCl. Elution fractions were separated by 12% SDS-PAGE. All the fractions containing the target proteins were concentrated using Amicon Ultra-4 Centrifugal filters (Merck Millipore) with 10 KDa membrane cut off, up to 10 mg.ml^{-1}. Measure of protein concentration was determined by

absorbance at 280 nm and using the extinction coefficient of 61800 $M^{-1}.cm^{-1}$ for PhoX and 64400 $M^{-1}.cm^{-1}$ for PstS, obtained with Protparam bioinformatic tool.

Western blot analysis was performed to confirm the expression of PhoX and PstS. An equivalent amount of 20 µg of each protein were dissolved in sample buffer, boiled for 5 min, and loaded onto a 15% SDS-PAGE gel at 100V for 1h. Separated proteins were then blotted onto polyvinylidene difluoride membranes (Amersham Pharmacia Biotech) using a transblot system (Bio-Rad) at 100V for 1h. The membrane was blocked for 1h in Tris-buffered saline (TBS) containing 5% non-fat dry milk, then it is washed three times with TBS containing 0.01% Tween 20 (TBST) for 15 min, and incubated overnight at 4°C with specific primary antibodies specific to the His-tag, Mouse Anti-His-Tag IgG1 Monoclonal Antibody (1:2500, R&D System- #MAB050). The next day, the membrane was washed three times with TBST and incubated for 1h in TBST with alkaline phosphatase conjugated antibody (secondary antibody), then it was washed three more times with TBST, and incubate with NBT/BCIP solution (Nitrotetrazolium Blue Chloride/ 5-bromo-4-chloro-3'-indolyl phosphate p-toluidine) at room temperature until desired color is achieved, finally membrane was washed 4 times with TBST.

3.1.3 Biophysical analyses

3.1.3.1 Denaturation and refolding of PstS and PhoX for binding assays

To eliminate potentially bound compounds that would affect the biophysical assays, PhoX and PstS were dialyzed in denaturation buffer (10 mM Tris-HCl, 6 M guanidine hydrochloride, pH 8.0) using a high purity guanidine hydrochloride (~99%, LGC Biotecnologia). Proteins were diluted to 20µM and dialyzed with denaturation buffer for 16h at 4°C in Dialysis Membranes (MWCO 10 kDa, Spectrapor). Then proteins were diluted to 10 µM and refolded by four consecutives dialyses of 12h at 4°C each one using 10 mM Tris-HCl pH 7.5 buffer and 10 mM NaCl. Aggregated proteins corresponding to misfolded species were removed by filtration using 0.22 µm cut-off filters.

3.1.3.2 Circular dichroism spectroscopy

Circular dichroism (CD) spectra of PhoX and PstS were recorded using a JASCO-815 spectrometer equipped with a Peltier temperature control system. Spectra of 1µM of

proteins (0.035mg/mL and 0.038mg/mL for PhoX and PstS, respectively) in 10 mM Tris-HCl pH 7.5 buffer and 5 mM NaCl, and in 5 mm path length quartz were recorded between 195 nm and 270 nm, at 25°C (298 K). The evaluated samples were native proteins and renatured proteins in presence of 1.5 μM or 30μM of Pi. The far-UV CD spectra were collected with 3 consecutive scans for sample and the average spectra were corrected by subtracting the buffer and ligand contribution. The secondary structure was estimated using the online tool, CAPITO- CD Analysis and Plotting Tool (https://capito.uni-jena.de/).

3.1.3.3 Small-angle X-ray scattering data collection and processing

Small-angle X-ray scattering (SAXS) experiments were performed in collaboration with Dr. Andre Sehnem and Prof. Cristiano Oliveira from the Institute of Physics at University of Sao Paulo (USP, São Paulo, Brazil), and with Dra. Aline Sampaio Cremonesi from the University of São Francisco (Jundiaí, São Paulo, Brazil). Measurements were performed on the Bruker-NANOSTAR SAXS equipment with a VÅNTEC-2000 detector. SAXS data were recorded on samples of PhoX and PstS, in their native or refolding states with or without Pi, in a molar ratio of 1:0.5 (protein: ligand). All solutions were measured at a final protein concentration of 2 mg.ml^{-1}, at 25°C in 10 mM Tris-HCl pH 7.5 buffer and 100 mM NaCl. Six independent measurements of 1800 s of exposure time, at least, were performed for each sample. The samples were filled in 1.5 mm thick cylindrical quartz-glass capillaries supported by homemade stainless-steel cases, and the scattering was detected using Pilatus image. The 2D isotropic scattering images were submitted to azimuthal integration, resulting in the 1D-SAXS intensity profiles, which are as a function of the scattering vector momentum (q) defined by $q=4\pi sin(\theta)/\lambda$, where 2θ is the scattering angle. Subsequently, data treatment of 1D SAXS profiles were performed with the subtraction of the buffer for each sample. SAXS profile of the samples were converted to an absolute scale using values of SAXS experiments from empty cell and water in the same conditions. The entire data were processed using the SUPERSAXS software package (Oliveira and Pedersen, unpublished). Then, the SAXS data profiles were fitted using the PRIMUS program, where the pair distribution function [p(r)] were calculated. The maximum diameter (Dmax) and the radius of gyration (Rg) were derived from the

results of the parameters p(r) and GNOM. Finally, the surface of the 3D structures was created with DAMIN, DAMAVER and SUBCOMB programs.

3.1.3.4 Intrinsic tryptophan fluorescence spectroscopy

Tryptophan fluorescence spectroscopy (TFS) experiments were performed in collaboration with Dr. Gabriel Vignoli from the University of Sao Paulo (USP, Sao Paulo, Brazil) at the USP Physics Institute. The Varian Cary Eclipse fluorimeter (Varian Inc, Santa Clara, CA) and a quartz cuvette of 3 ml with 4 mm of optical pathway were used. To measure PhoX-Pi and PstS-Pi interactions, the excitation beam light was set to 295 nm and the temperature was maintained constant at 25°C. Proteins and ligand were prepared using the same buffer 10 mM Tris-HCl pH 7.5 and 100 mM NaCl, the resulting solutions were 1 ml of proteins at 10 µM, either PhoX or PstS, and a stock solution of Pi at 1 mM. Proteins were homogenized for 10 minutes before being titrated with Pi. PhoX and PstS were tested independently with increasing concentrations of ligand, for PhoX from 0 to 110 µM of Pi, and for PstS from 0 to 158 µM of Pi. Data from the experiments were analysed in Origin 8 software and no inner filter correction was necessary, as absorbance values at this wavelength were smaller than 0.05. Due to the fact that both proteins have more than one tryptophane, the fluorescence spectra were described with the Stern-Volmer equation: $F_0/F = 1 + K_{SV}[Q]$, where K_{SV} is the product of the natural radiative lifetime (the lifetime in the absence of quencher) known as the Stern-Volmer constant, [Q] is the quencher or ligand concentration and F_0/F is the resulting ratio of emission intensities (Lakowicz, 2006).

3.1.4 Bioinformatic analyses

3.1.4.1 PstS modeling

The three-dimensional model of PstS, and the molecular dynamics of PhoX and PstS (proteins from *X. citri*) were carried out in collaboration with Prof. Ignacio General from the National University of General San Martin (UNSAM, Buenos Aires, Argentina) and B.S. Georcki Ropon from the National University San Antonio Abad del Cusco (UNSAAC, Cusco, Peru). The model of *X. citri* PstS was obtained using the structural coordinates of *Yersinia pestis* PstS (PDB code 2z22) as template (sharing 52.26% of amino acid

sequence identity), Clustal Omega and Modeller programs were used to align the amino acids sequences and produce 1000 models by homology, respectively. Generated models were analyzed, and it was selected the model with 0.091 of RMSD and no residues in regions generously allowed or disallowed.

3.1.4.2 Molecular dynamic of PhoX and PstS

In silico molecular dynamics of PstS and PhoX were performed using the PstS model and the crystallographic structure of PhoX (PDB identification code 5I84). The Amber18 program was used to perform the simulations of four systems: apoprotein PstS, apoprotein PhoX and the complexes PstS-PO_4^{3-} and PhoX-PO_4^{3-}. Simulations in each system were repeated at least two times and had a duration of 1.5 or 1.0 µs. Root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF) were determined using the software Groningen machine for chemical simulation (GROMACS) and virtual molecular dynamics (VMD). At the end, it was calculated the free energy difference between proteins with and without phosphate.

3.1.5 Crystallization assay of PstS

PstS was expressed and purified as previously described using IMAC and SEC purification and three dialysis of 8 hours with 10 mM Tris-HCl pH 7.5 buffer and 10 mM NaCl. First screening was performed with recombinant PstS at 10 mg.ml^{-1} in 10 mM Tris-HCl pH 7.5 buffer and 10 mM NaCl. A total of 98 buffers from the crystallization kits, Crystal Screen 2™ HR2-110 and HR2-112 (Hampton Research), were used using the hanging-drop vapour-diffusion method, with a molar ratio of 1:1 (protein: buffer) and incubating at 18°C. Then, the second screening were done with recombinant PstS at 23.5 mg.ml^{-1} in 10 mM Tris-HCl pH 7.5 buffer and 10 mM NaCl. Three crystallization kits were used: JBscreen Classic 1 to 4, JBscreen Basic 1 to 4 and Pi-minimal screen (Jena bioscience), and the sitting-drop vapour-diffusion method was applied with the same molar ratio and storage temperature. Protein crystals were colored by a solution of 0.05% (w/v) of methylene blue.

3.2 Chapter II

3.2.1 Bacterial strains and plasmids

The strains used were *Escherichia coli* BL21(DE3) and *Xanthomonas citri* subsp. *citri* 306, the plasmids used were pOP3BP, pUCP30T, pKD46 and pKD13.

3.2.2 Bioinformatic analyses of chemotactic proteins and its relation with Pst system

To identify the putative chemotactic proteins in *X. citri*, a basic local alignment search (Blastp of *X. citri*) was performed using the *P. aeruginosa* CtpH (PA2561) and CtpL (PA4844) amino acid sequences as queries. As a result of the previous analyses Tsr1 (KEGG identifier: XAC1896) and Tsr2 (KEGG identifier: XAC0611) were selected to be studied. Then, the amino acids sequences of these proteins were analyzed with the bioinformatics servers "DAS"–Transmembrane Prediction (available in https://tmdas.bioinfo.se/DAS/index.html) and TOPCONS (available in http://topcons.cbr.su.se/). The analysis allowed the recognition of the Ligand-Binding Domains (LBD), which are the part of the sensory domain located in the periplasm.

3.2.3 Cloning, expression and purification of chemotactic proteins

Nucleotide sequences encoding Tsr1$_{LBD}$ (residues from 6 to 131) and Tsr2$_{LBD}$ (residues from 2 to 129) were amplified from *Xanthomonas citri* subsp*. citri 306 (X. citri)* genome by polymerase chain reaction (PCR) using the following primers: 5'-AAA**GGATCC**GGCGAATTCCGCGTCTACGAAG-3' and 5'-AAA**AAGCTT**TTACTTGAAC GAGCTGGCGCTG-3' for Tsr1$_{LBD}$, and 5'-AAA**GGATCC**CAGCACTATGCCGATAAGGT CAAAG-3' and 5'-AAA**AAGCTT**TTAGCCCCACGCCTGTACTTG-3' for Tsr2$_{LBD}$. Both fragments were cloned into pOP3BP *Bam*HI and *Hin*dIII sites, and the proteins were produced from the resulting plasmids pOP3BP-Tsr1$_{LBD}$ and pOP3BP-Tsr2$_{LBD}$ and under the control of T7 promotor/ lac operator **(Appendix 2A, 2B).** Plasmids were transformed in *E. coli* BL21(DE3) cells, and cultured in 4 liters of LB medium (10 g.l^{-1} of peptone, 10 g.l^{-1} of yeast extract and 5 g.l^{-1} of sodium chloride) supplemented with 100 μg.ml^{-1} of ampicillin at 37°C. When cells reach an OD$_{600}$ of 0.5, the expression of recombinant proteins were induced by adding 0.5 mM of IPTG for pOP3BP-Tsr1$_{LBD}$ and 0.1 mM of

IPTG for pOP3BP-Tsr2$_{LBD}$, keeping growing for 3h at 37°C and then the cells were collected by centrifugation. To purify Tsr1$_{LBD}$, cells were resuspended in the lysis buffer, 30 mM Tris-HCl buffer pH 7.5, 300 mM NaCl, 10 mM imidazole and 5% of glycerol, and sonicated on ice. After centrifugation, the supernatant was applied to an IMAC column (HisTrap Chelating HP, GE Healthcare Life Sciences) equilibrated with Buffer B (Tris-HCl 30 mM pH 7.5, NaCl 300 mM and Imidazole 10 mM), Tsr1$_{LBD}$ was eluted using an imidazole concentration between 50- and 500- mM in Buffer B. For Tsr2$_{LBD}$ purification, cells were lysed by sonication on ice using Tris-HCl 50 mM pH 7.5 buffer, Glycerine 10%, NaCl 100 mM, Triton x-100 0.0003%, Tween 20 0.03% and sucrose 20%. Then, cells were harvested by centrifugation and Tsr2$_{LBD}$, were separated from the supernatant by IMAC using Buffer C (Tris-HCl 50 mM pH 7.5, NaCl 100 mM and Imidazole 20 mM) and a gradient imidazole concentration in a range from 100- to -500 mM. Tsr1$_{LBD}$ and Tsr2$_{LBD}$ domains were further purified by SEC in Hi Load 16/60 200 Superdex column (GE Healthcare Life Science) equilibrated with 30 mM Tris-HCl pH 7.5 buffer and 100 mM NaCl. Fractions containing the target domains were concentrated in Amicon Ultra-4 Centrifugal filters (Merck Millipore) with 10 KDa membrane cut off. Typically, 1 mg of purified Tsr1$_{LBD}$ and Tsr2$_{LBD}$ were obtained per litre of bacterial culture. Subsequently, both proteins were treated overnight with 3C protease to cleavage the GB1 fusion protein and were submitted again to SEC to separate the domain of interest. At the end, 1 ml of 0.3 mg were obtained of each domain.

3.2.4 Vector design to knockout phoX gene from X. citri

The pUCP30T plasmid is a suicide vector (cannot replicate) in *Sinorhizobium meliloti* (*S. meliloti*). To test if it could also be a suicide vector in *X. citri*, a colony of *X. citri* was grown in 5 ml of 2XTY medium (16 g.l^{-1} of tryptone, 10 g.l^{-1} of yeast extract and 5 g.l^{-1} of sodium chloride) overnight at 30°C, then 1 ml was inoculated in 50 ml of 2XTY medium at the same temperature up to OD600 of 0.3. Cells were collected by centrifugation; pellet was washed three time with ice-cold sterile water and the final pellet was resuspended in the residual water, 50 µl of these cells were mixture with the 3 µl (200 ng/ µl) of plasmid pUCP30T. Transformation was done by electroporation using the BIO-RAD system, cells plus plasmid were placed into a 0.2 cm cuvette (BIO-RAD) and electroporated at 2KV,

25uF and 200V. Cells were recovered in 1 ml of 2xTY and incubated for 2 hours at 30°C, thereupon cells were centrifugated and the pellet was resuspended in 100 µl of 2xTY medium and plated on LB agar (10 g.l^{-1} of peptone, 10 g.l^{-1} of yeast extract, 5 g.l^{-1} of sodium chloride and 15 g.l^{-1} of agar) supplemented with gentamycin (40 µg.ml^{-1}) and incubated at 30 °C for 3-4 days. *No colonies were observed on the plates.*

Due to the previous result the pUCP30T plasmid was used to do a one-step knockout of *pho*X gene in *X. citri*. A fragment of approximately 2000 pb, containing the *phoX* gene flanked by i500 pb upstream and downstream, was amplified from *X. citri* by PCR using the primers: 5'-CCC**TCTAGA**AACCTCACCGGCGAAACCT-3' and 5'-CCC**AAGCTT**AAATTTCAGCACCCTTGCTGTCG-3'. The nucleotide fragment was cloned into pUCP30T originating the plasmid pUCP30T-up-phoX-dw**,** and its nucleotide composition was confirmed by sequencing using the commercial primers M13 forward and reverse (ThermoFisher Scientific). Afterwards, it was used an *E. coli* BW25113 strain (*E. coli* BW25113) carrying a Red helper plasmid called pKD46 **(Appendix 3A).** This plasmid encodes the Exo, Beta and Gam proteins that constitute the Lambda Red recombination system, which promotes an enhanced rate of recombination between a plasmid and a linear DNA. Furthermore, pKD46 has an ampicillin resistance cassette, that is under control of *paraB* promotor/ L-arabinose operator, and it is thermosensitive at 37°C. The idea is introducing a linear DNA into *E. coli* BW25113 containing pKD46 and pUCP30T-up-phoX-dw plasmids to delete the *phoX* gene in pUCP30T-up-phoX-dw plasmid. For this purpose, the linear DNA up-km-dw, was amplified from the plasmid pKD13 **(Appendix 3B)**, using the primers: 5'-AATGGCACCACGGCGGGACGTCCTCAT CCCTTCGCTACAGGAGCTGTTCC**GTGTAGGCTGGAGCTGCTTC**-3' and 5'-AGCATG GCTTGCCCCACCAGGCCGCTGCAGCAGAGGCCTGCGCCGTACTC**ATTCCGGGGA TCCGTCGAC**-3'. This linear carries a kanamycin resistance gene flanked in 5' with a FRT region and 50 pb upstream the *phoX* gene, and in 3' with a FRT region and 50 pb downstream the *phoX* gene. Subsequently, transformants carrying the plasmids pKD46 and pUCP30T-up-phoX-dw were grown in 5 ml SOB medium (BIO BASIC) with 100 ug.ml^{-1} of ampicillin and 1 mM of L-arabinose at 30°C to an OD600 of 0.6 and then made electrocompetent by concentrating and washing three times with ice-cold 10% glycerol. Electroporation was done pouring the 50 µL of cells and 100 ng of the linear DNA up-km-

dw into a chilled 0.15 cm cuvette and setting at 1.75 kV, 25 uF and 200 ohms. Shocked cells were added to 1 ml of SOC medium (Tryptone 2%, Yeast extract 0.5%, KCl 2.5 mM, MgCl$_2$ 10 mM, MgSO$_4$ 10 mM and Glucose 20mM), incubated 1 h at 37°C, and then one-half was spread onto LB agar with 100 ug.ml^{-1} of kanamycin and incubated overnight at 37°C, then colonies were tested for ampicillin sensitivity to test for loss of the helper plasmid. Finally, the nucleotide sequence of the resulting knockout plasmid pUCP30T-up-km-dw was verified by sequencing. Then the pUCP30T-up-km-dw plasmid was electroporated to *X. citri* trying to get *X. citri* strains with *pho*X gene deletion.

An additional advantage of the designed plasmid is that it could be used as a template to amplify the sequence up-km-dw for subcloning into known suicide vectors of *X. citri* such as pNPTS138.

4 RESULTS AND DISCUSSION

4.1 Chapter I

4.1.1   PhoX and PstS were produced as soluble and stable proteins

The recombinant proteins PstS and PhoX from *X. citri* were overexpressed in *E. coli* BL21(DE3) cells with the expected theoretical molecular weight of 37.94 KDa and 35.46 KDa, respectively (PEGOS *et al.*, 2014 and PEGOS *et al.*, 2017). After the bacterial cell lysis with 50 mM Phosphate buffer pH 7.2, 100 mM NaCl, 20 mM imidazole and 5% of glycerol, most of the proteins were in the soluble fraction **(Figure 16A, 16B)**. Proteins were purified with no contaminants by IMAC in 50 mM Phosphate buffer pH 7.2, 100 mM NaCl, and started to elute with different imidazole concentrations: 68 mM for PstS and 36 mM for PhoX **(Figure 16C, 16D)**. A second step of purification by SEC using 10 mM Tris-HCl pH 7.5 buffer and 10 mM NaCl, was used to eliminate the imidazole and exchange the buffer for subsequent analysis **(Figure 16E, 16F)**. At the end, PhoX and PstS were produced with an efficiency of 18 mg.l^{-1} and 12 mg.l^{-1}, respectively.

To confirm the production of the recombinant proteins, the purified samples were submitted to Western blot analysis using the Mouse Anti-His-Tag Monoclonal Antibody (R&D System) **(Figure 16G)**. The recombinant proteins were recognized by the antibodies in all the corresponding fractions.
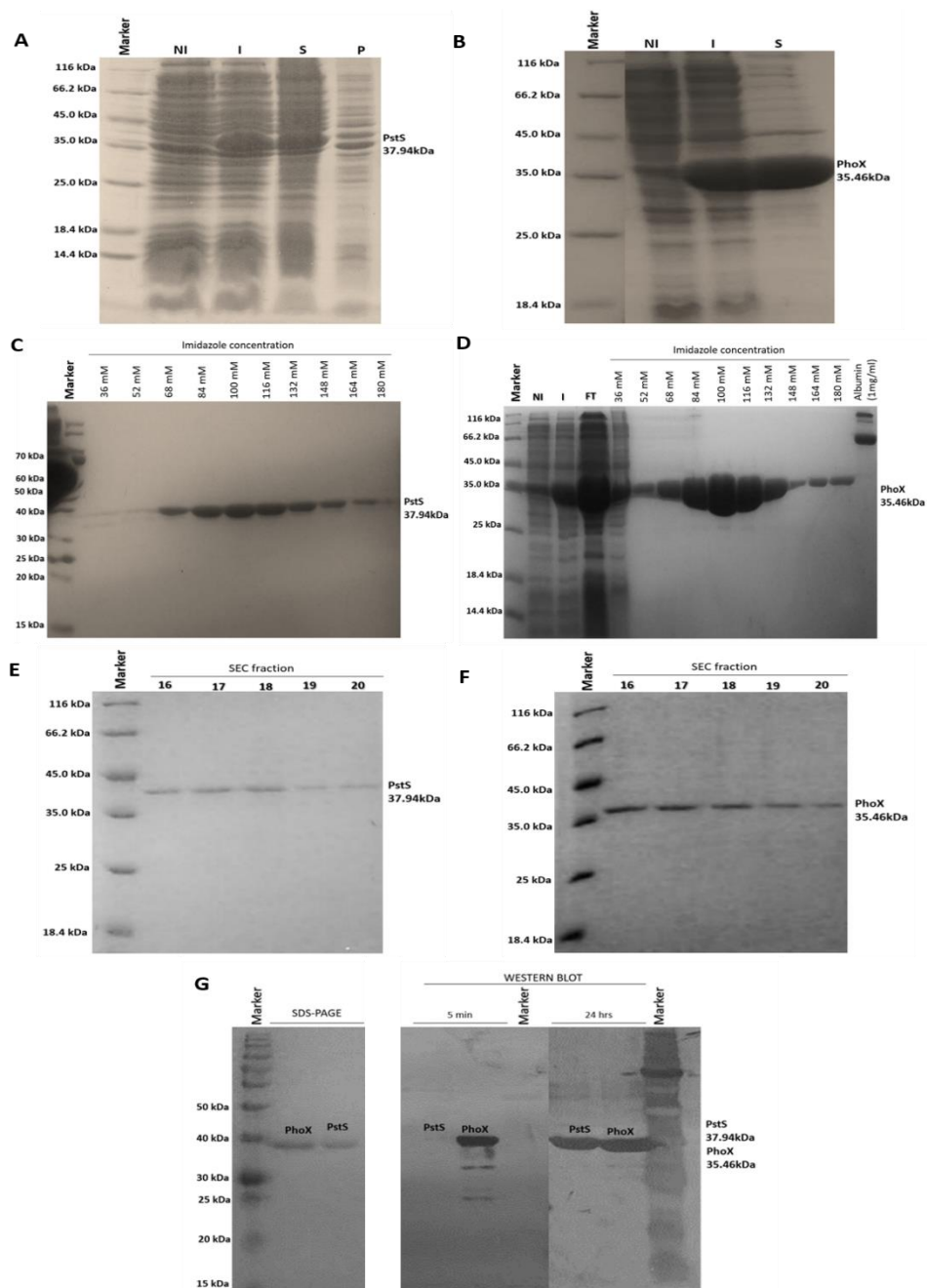
Figure 16. Purification and Western blot of _X. citri_ PstS and PhoX. **A** and **B.** SDS-PAGE 12% showing the overexpression of PstS (37.94 KDa) and PhoX (35.46 KDa), respectively, in BL21(DE3) strain using 0.1 mM of IPTG and 4 hours of induction at 30°C. In both cases, most of the proteins are in the soluble fraction. NI: Not induced, I: Induced, S: Supernatant fraction, P: Pellet fraction. **C** and **D.** SDS-PAGE of samples coming from purified PstS and PhoX by Immobilized Nickel-Affinity Chromatography (IMAC) and in different concentration of imidazole. Both proteins were purified successfully, PstS starts eluting at 68 mM of imidazole and PhoX at 36 mM of imidazole. **E** and **F.** SDS-PAGE showing samples of PstS and PhoX proteins separated by Size Exclusion Chromatography (SEC) using 10 mM Tris-HCl pH 7.5 buffer and 10 mM NaCl. Both proteins were eluted in the same fractions (from 16 to 20) because of their similar molecular weights. **G.** Purified PstS and PhoX showed in SDS-PAGE (left) and detected by Western blot using the Mouse Anti-His (1:2500, R&D System) (right). PhoX was stained with NBT/BCIP solution after 5 minutes, while PstS needed 24 hours.

4.1.2 Biophysical analyses

4.1.2.1 Renatured PhoX and PstS proteins conserve their folding

To evaluate the influence of phosphate in the secondary structure content of the proteins and to calculated their affinities for the anion, biophysical experiments were performed. Since the proteins could already be purified with phosphate in the ligand-binding site, two samples of each protein were produced, the *native* and the *renatured*. The *native* PstS (N-PstS) or (N-PhoX) were obtained after the IMAC, SEC and steps of dialysis in 10 mM Tris-HCl pH 7.5 buffer and 10 mM NaCl, while the *renatured* proteins (R-PstS and R-PhoX) were native proteins denatured with guanidine and renatured by dialysis in the same buffer than for native proteins.

Circular Dichroism (CD) assays were performed to compare the folding and the secondary structural content of the proteins, as well as the efficiency of the renaturation process. CD is a spectroscopy technique that gives an estimation of the characteristic folding of a protein with percentage of secondary structures such as alpha-helices, β-strands and irregular structures as loops (GREENFIELD, N. 2006; MICSONAI *et al.* 2015). Firstly, the CD profile of the native and renatured proteins showed that PstS and PhoX were successfully refolded, presenting a typical and expected alpha-beta folding (mix of two negative peaks at 222 nm and 208 nm evidencing α-helix topology, and a negative peak at 218 nm and a positive peak at 195 nm usual of β-strands) characteristic of substrate-binding proteins that belong to ABC transporters (GREENFIELD, N. 2006; BERNTSSON *et al.,* 2010; SCHEEPERS *et al.*, 2016; BARTUAL *et al.,* 2017). The comparison between PstS and PhoX CD reveals that they share similar profiles. In the same way, after the denaturation-renaturation process both proteins suffered a slight loss of secondary structure content. N-PstS and R-PstS behavior shows the presence of a higher content of non-structured or irregular regions in comparison to PhoX samples, suggesting that PstS is less stable than PhoX **(Figure 17)**.

A

Native and renatured PstS

	Prediction based on aa sequence	Native PstS	Renatured PstS
Helix	0.35	0.28	0.26
B-strand	0.31	0.41	0.39
Irregular	0.34	0.31	0.35

B

Native and renatured PhoX

	Prediction based on aa sequence	PhoX Structure	Native PhoX	Renatured PhoX
Helix	0.43	0.41	0.27	0.26
B-strand	0.35	0.16	0.42	0.41
Irregular	0.23	0.43	0.31	0.33
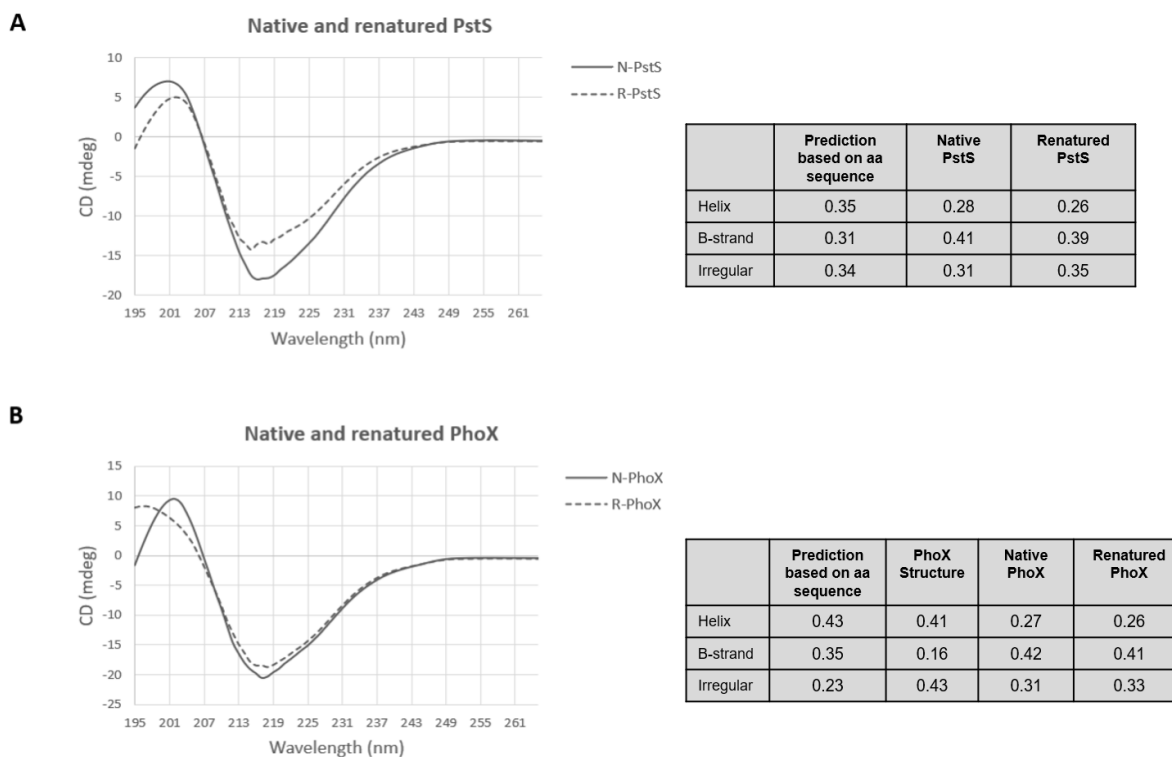
Figure 17. Folding and secondary structure content of PstS and PhoX evaluated by Circular Dichroism (CD) before and after the renaturation process. Purified proteins were submitted to denaturation process with guanidine 6 M and then refolded. **A.** CD spectra of the native and renatured PstS (N- and R-, respectively) revealed that after refolding the protein achieved a similar folding, but had loss of alpha-helix and beta-sheet content, as evidenced by the increasing of the non-structured content. **B.** Secondary structure prediction of PhoX (N- and R-) revealing that renatured PhoX profile was very similar the native with a slight loss of secondary structure. The quantification of secondary structure content of PstS and PhoX in both conditions proteins is presented in the Tables. The PhoX results also compares the secondary structure content from the PhoX crystal structure bound to phosphate (PDB 5I84).

4.1.2.2 Structural changes of PstS and PhoX in presence of phosphate

In periplasmic-binding proteins as PstS and PhoX, the ligand-binding site is in the cleft between the N and C domains. Upon ligand binding and residues interactions, it is expected the approximation of the domains, which can be translated in structural changes, sometimes identified in CD analysis. The structural changes induced by phosphate adding in the renatured PstS and PhoX samples, two different concentrations of the ligand were used in the CD assays. Both proteins showed a slight gain of secondary structure content upon phosphate addition that might suggest the domains approximation. According the calculations of the CAPITO program, the interaction of PhoX with phosphate, promotes a small gain in the alpha-helix content. In PstS the gain is distributed between alpha-helices

and beta-sheets, indicating a higher flexibility of the secondary structures in this protein than in PhoX, as evidenced in **Figure 18**. These results are different to the previously showed by PEGOS and collaborators (2017) that did not identified significant differences in the proteins after phosphate addition. Since in that work the proteins were not submitted to denaturation/renaturation processes, our results evidence the importance of denaturation for obtaining of a free-ligand protein. The structural changes induced by ligand binding were described for other SBPs and not always evidenced gain of secondary structures. The *E. coli* ribose-binding protein RbsB (Ribose-binding protein) lost secondary structures in the presence of ribose (TAVARES *et al.*, 2019) but the *X. citri* MalE (Maltose-binding protein) showed a significant gain of secondary structure after maltose binding (BALAN *et al.*, 2005).

CD is a good assay to evaluate structural changes in general, but it gives only estimations of structural content (MICSONAI *et al.*, 2015) and should be complemented with other techniques.

A

PstS in presence of PO₄⁻

	Prediction based on aa sequence	Renatured PstS (1μM)	Renatured PstS + 1.5 μM PO₄⁻	Renatured PstS + 30 μM PO₄⁻
Helix	0.35	0.26	0.26	0.27
B-strand	0.31	0.39	0.42	0.41
Irregular	0.34	0.35	0.32	0.32

B

PhoX in presence of PO₄⁻

	Prediction based on aa sequence	Renatured PhoX (1μM)	Renatured PhoX + 1.5 μM PO₄⁻	Renatured PhoX + 30 μM PO₄⁻
Helix	0.35	0.26	0.34	0.32
B-strand	0.31	0.41	0.36	0.39
Irregular	0.34	0.33	0.30	0.29

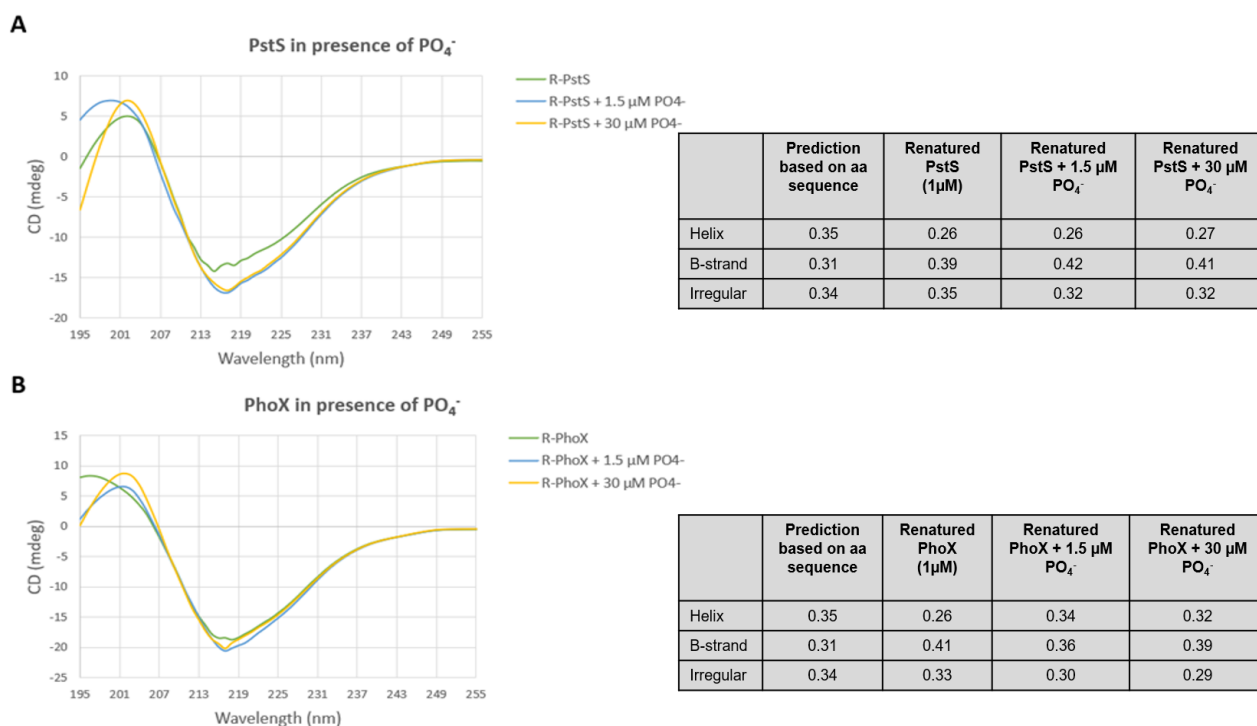**Figure 18. Phosphate influence in the secondary structure content of PstS and PhoX.** CD Profiles of R-PstS **(A)** and R-PhoX **(B)** in presence of phosphate showing a slight gain of alpha-helices and beta-sheets in PstS and notorious gain of alpha-helices in PhoX. Differences can suggest a general flexibility of the secondary structures in PstS in opposition to PhoX.

4.1.2.3 SAXS analyses of PstS and PhoX before and after phosphate binding

The structural behavior of PstS and PhoX in presence and absence of phosphate also was evaluated by SAXS experiments. Unfortunately, the SAXS data collected from PstS samples could not be processed due to possible protein aggregation. On the other hand, the scattering patterns from PhoX apo- and bound- proteins are presented in **Figure 19**. The one-dimensional curves of the two samples indicate that PhoX suffers changes in its tridimensional conformation upon ligand binding, as evidenced in the CD analysis. Interestingly, according to the scattering patterns of PhoX **(Figure 19A)**, the apoprotein is more stable than the bound protein, which was not expected because PBPs tend to stabilize the protein after the interaction protein-ligand (SHEVCHUK *et al.*, 2017). The Kratky plots show that part of PhoX proteins (bound and unbound) have a high flexible region more evident in the bound protein **(Figure 19B).** Moreover, Primus program was used to generate the graph of P(r), the distance distribution, and calculate the Radius of Gyration (Rg). The results indicate that PhoX shows a maximum size around 90 Å and 85 Å for the apo- and apo- proteins, respectively. Accordingly to the P(r) curves, the geometrical shape of the apoprotein is oblate ellipsoid while the bound protein is well represented as two-domain protein **(Figure 19C)**.

Protein envelopes of the both states of PhoX were created by DAMIN, DAMAVER and SUBCOM programs and revealed interesting differences between PhoX or PhoX-PO_4^{3-}. PhoX was characterized as a single envelope in contrast to the PhoX-PO_4^{3-}, which presented an envelope like a fish representation, with a tail (small domain of 31.5 Å) and the body (approx. 59.5 Å). To explain why both states of PhoX assume different shapes, the function of this protein must be analyzed. As mentioned, PhoX is a periplasmic-binding protein from an ABC transporter, and as such, presents two domains, which in the unbound state should be performing a movement of opening and closing, conferred by the hinge flexibility. Once the ligand achieves the ligand-binding pocket the domains get closer for ligand coordination and probably loose the flexibility. In addition, PhoX used in this experiment had the N-terminal His6-tag of 21 amino acids, which was not cleaved. Knowing that, we propose that the differences identified in the PhoX envelopes presented in this work, describe the flexibility of the tag, in the both apo- and bound- states. PhoX envelope in absence of phosphate is ellipsoid and suggests a protein of one domain,

which is acceptable if we think that the His6tag accompanies the movement of N-domain in synchronicity with the C-domain. In this case, the protein flexibility is generalized. On the other hand, the data for PhoX in the presence of phosphate show an envelope that looks like a fish where the tail is a small domain and the body a large one. If phosphate induces the domains closure and their loss of flexibility, it is supposed that they would move as a unique body connected to a flexible His6tag that would move individually. In fact, the data show that either Dmax and Rg values decrease in the presence of phosphate (Dmax of 90 Å to 85 Å and Rg of 29.03 ± 0.34 Å to 27.85 ± 0.351 Å) **(Figure 20)**. This result is in consonance with results obtained for the Glucose-Binding Protein (REINARTZ *et al.,* 2021), Maltodextrin-Binding Protein (MBP) (STOCKNER *et al.*, 2005) and Ribose-Binding Protein (RBP) (LI *et al.*, 2021), all of them showing a decreasing of Rg when bound ribose.

In accordance with the three-dimensional structure of PhoX resolved with phosphate in the ligand-binding pocket, the maximum distance (Dmax) between two opposite residues, located at the extremes of each domain, is around 61 Å (PEGOS *et al.,* 2014). Considering an additional N- terminal his-tag of 31.5 Å (assuming 1.5 Å between two residues, as previously reported (BERG *et al.,* 2002; CARTAILLER, J. and LUECKE, H.; 2004), these values are in accordance with the values of 90 Å and 85 Å calculated for the PhoX envelopes. This model that explains the envelopes can also justify the higher stability of the bound PhoX in the scattering plot and the results of flexibility observed in the Kratky plot.

Once the molar ratio of protein:ligand in the SAXS assay was 1:0.5, theoretically, a half quantity of protein should be bound to phosphate and the other half unbound. In the PhoX envelope in the presence of phosphate, we observed that the region corresponding to the C-domain shows a larger volume in comparison with the N-terminal, suggesting a higher flexibility of C-domain. In fact, C-domain contains more non-folded regions than N-domain, but we highlight an alpha-helix that might achieve high flexibility due its connection with non-structured loops **(Figure 21,** in magenta**).**
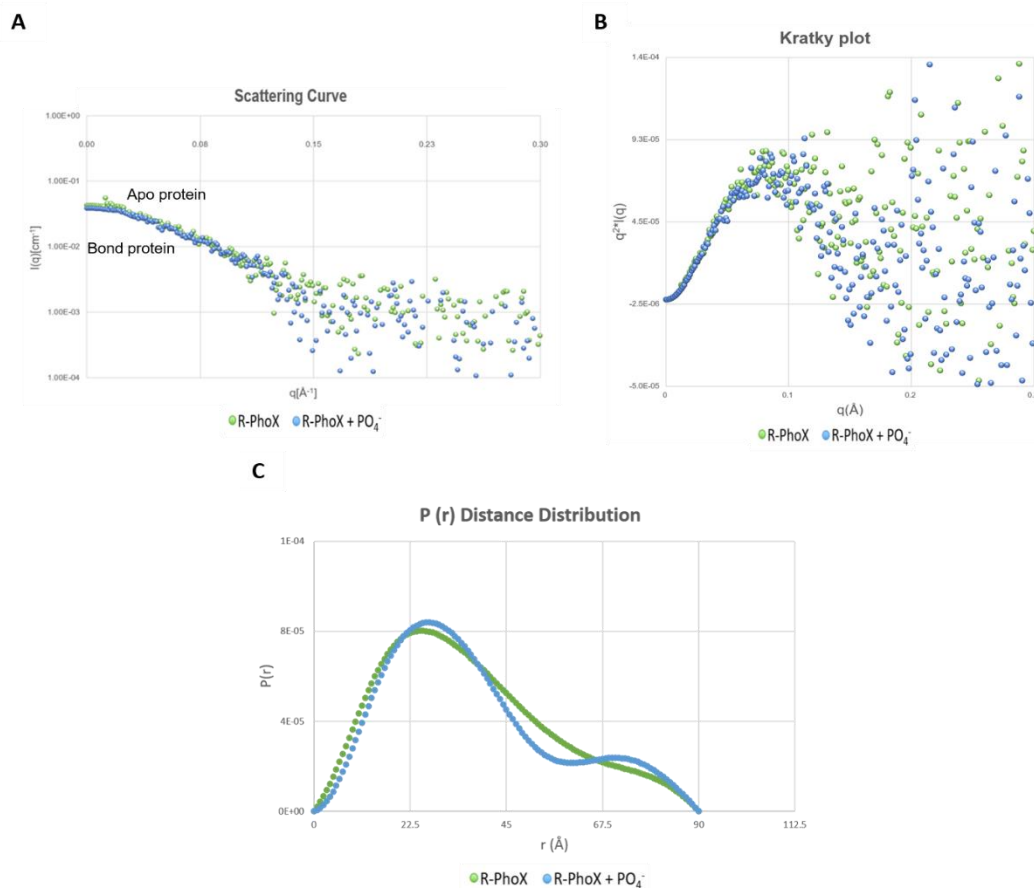
Figure 19. The scattering pattern curves, the Kratky plots and P(r) distances distribution of PhoX proteins obtained from SAXS experiments. A. The scattering patterns generated for PhoX and PhoX-PO_4^{3-} showing slight differences. **B.** Kratky plot of PhoX and PhoX-PO_4^{3-} showing high flexibility in absence and presence of phosphate. **C.** Experimental P(r) functions of PhoX and PhoX-PO_4^{3-} revealed Dmax around 90 Å, but PhoX in presence of the anion assume a multimeric form.
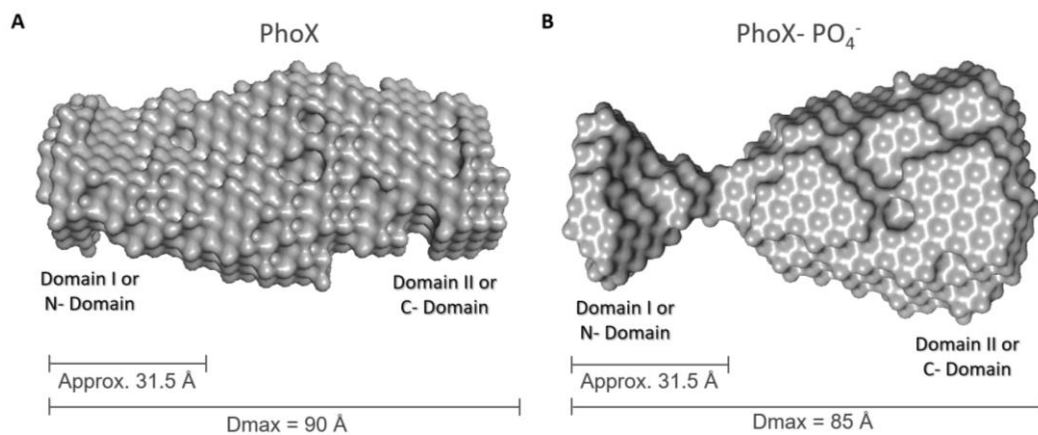
Figure 20. Comparison of PhoX envelopes in absence and presence of phosphate. A. PhoX apoprotein envelope suggests free movement of both domains and the his6tag in consonance. **B**. PhoX-PO_4^{3-} envelope shows a fish shape, with two defined regions, the his6tag N-terminal with 23 residues with a length of 31.5 Å (corresponding to the fish tail) and the closed domains (fish body).
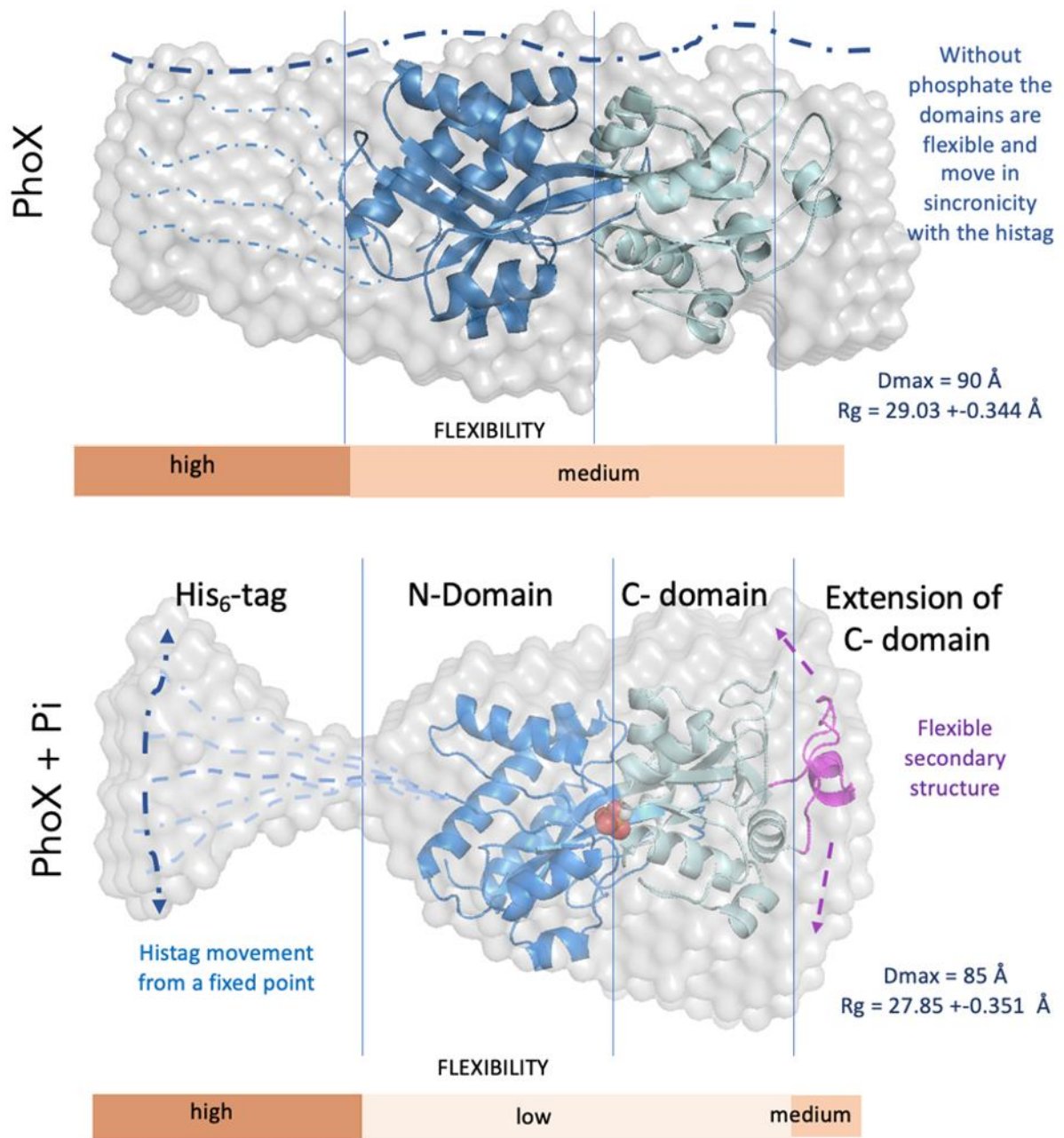
Figure 21. Model to explain the differences on PhoX domains organization in absence and presence of phosphate. When PhoX is unbound, the domains seem to have similar freedom of movement (flytrap), which are synchronized with the N-terminal histag. The resultant envelope describes one uniform protein. On the other hand, when PhoX binds phosphate (ligand shown as red spheres) it promotes the stability of both domains that close around the ligand. In this case, the his6tag is now connected to a stable structure and have freedom of movement from one unique point, generating the constriction in the envelope. The envelope enlargement in the region corresponding to C-domain (cyan color) is explained by the presence of a higher number of flexible secondary structure elements in this domain than in N-domain, like the flexible helix in magenta, which movement enlarge that part of the envelope. The arrows show the predicted movement of protein and tags.

58

4.1.2.4 Tryptophan Fluorescence Spectroscopy (TFS) to determinate the PstS and PhoX affinities for phosphate.

Tryptophan (Trp or W) emits intrinsic fluorescence that can be measured by spectroscopy (TFS) after excitation at 295 nm, and emission at ~350 nm (ERIKSSON *et al.*, 1993; GHISAIDOOBE AND CHUNG, 2014; SINDREWICZ *et al.*, 2019). In addition, this residue has at least two different fluorescence lifetimes (~0.5 and ~3.1 ns) that justify its high sensitivity to the environment (GHISAIDOOBE AND CHUNG, 2014).

If a Trp residue in a protein is directly involved in the ligand binding, or if its environment is altered upon ligand interaction, it is possible to measure the variation of the fluorescence intensity in this environment (LAKOWICKS *et al.*, 2006, GHISAIDOOBE AND CHUNG, 2014; SINDREWICZ *et al.*, 2019). The fluorescence measuring of a sample in the presence of increasing concentrations of a specific ligand can be used for calculation of the dissociation constants and the ligand affinities. In the case of proteins that contain several tryptophan residues, the complex Stern-Volmer plots can be used to calculate the Kd (LAKOWICKS *et al.*, 2006).

This is exactly the case of *X. citri* PstS and PhoX proteins, which have 9 and 8 tryptophans, respectively. In accordance with the resolved 3D structure of PhoX bound to phosphate (PDB identification code 5I84) (PEGOS *et al.*, 2014), there are four tryptophans in the Domain I(W43, W257, W305 and W334) and four in the Domain II(W135, W180, W192, and W276) **(Figure 22A)**. The amino acid sequence alignment of PhoX with PstS shows that PstS has 9 tryptophan, 8 in the same position than PhoX (Domain I: W67, W281, W329, W358 and Domain II: W159, W204, W216, W300) and an additional (W264) in the Domain II **(Figure 22B)**.
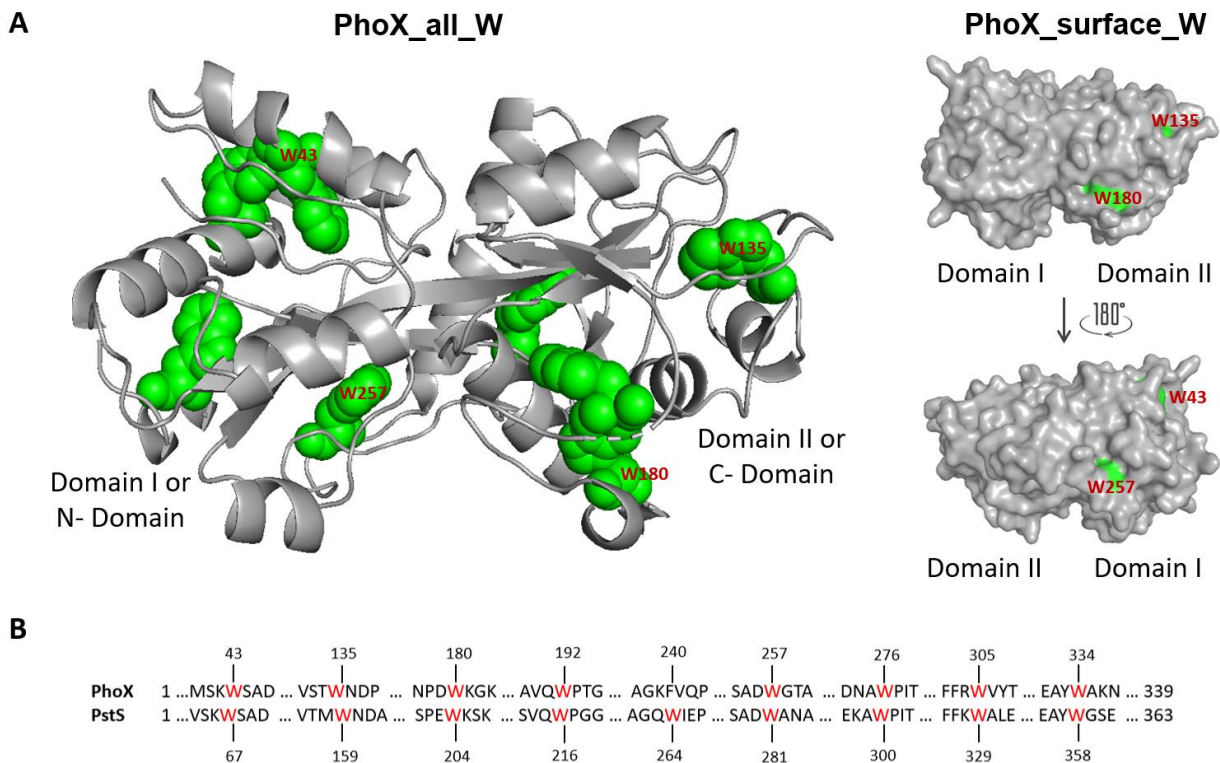
Figure 22. Tryptophans distribution in *X. citri* PhoX. **A.** On the left, PhoX structure showed in cartoon with a front view of the phosphate-binding pocket. The tryptophan (W) are show in green spheres, four in Domain I and four in domain II. On the right, PhoX surface with a front view of the phosphate pocket in the up, and back view in the down. In both cases, the contribution of tryptophan emission can be observed on the surface of the protein. **B.** Amino acid sequence alignment of PhoX and PstS highlighting the 8 conserved tryptophans and the additional residue W264 of PstS.

Most PeBPs inherently yield changes in the intrinsic fluorescence upon binding to their cognate ligands (EDWARDS K., 2021). This applies to *X. citri* PhoX, where TFS technique detected the quenching of tryptophan in the presence of phosphate (PO_4^{3-}), but not to PstS. The maximum Trp fluorescence of PhoX during titration occurred at 325 nm **(Figure 23A, 23B)**. As the presence of phosphate suppresses PhoX fluorescence, the Stern-Volmer equation (Equation 1) was applied to obtain more information about the PhoX-PO_4^{3-} interaction.

$$\text{Equation 1.} \quad \frac{F_0}{F} = 1 + K_{SV}[Q]$$

where K_{SV} is the Stern-Volmer constant, [Q] is the ligand concentration and F_0/F is the resulting ratio of emission intensities. In case that the Stern-Volmer plot (F_0/F Vs. [Q]) does not show a linear behaviour and shows a curve with two concavities downwards **(Figure 23C)**, it means that there are two different populations of fluorophores, one accessible to the suppressor and another with less accessibility. In this case we had to use the modified Stern-Volmer equation (Equation 2).

$$\text{Equation 2.} \quad F = \frac{F_{0a}}{1 + K_{SV}[Q]} + F_{0b}$$

where F is the total fluorescence in the absence of ligand, F_{0a} is the fluorescence of the accessible fluorophore, F_{0b} is the fluorescence of the buried fluorophore, K_{SV} is the Stern-Volmer quenching constant of the accessible fraction and [Q] is the ligand concentration. Additionally, we used the Equation 3 to calculate the fraction of the initial fluorescence that is accessible to quencher, *f*a the **(Figure 23D)**.

$$\text{Equation 3.} \quad f\text{a} = \frac{F_{0a}}{F_{0b} + F_{0a}}$$

The modified Stern-Volmer plot ($F_0/F-F_0$ Vs. 1/[Q]) shows the range of ligand concentrations used which strongly suggests that the product is a non-fluorescent complex indicating a static suppression. As a static suppression the Stern-Volmer constant is equal to the Equilibrium binding constant (K_d) (Equation 4).

$$\text{Equation 4.} \quad K_{SV} = \frac{1}{[K_d]}$$

The modified Stern-Volmer plot revealed that the PhoX K_d for phosphate is 6.1 µM **(Figure 23C),** which is in accordance with the literature that showed Kd for other SBPs in the micromolar range. Examples of that are the *E. coli* Maltodextrin-Binding Protein (K_d for maltose of 3.5 µM) (SPURLINO *et al.*, 1991), the Sulfate-Binding Protein (K_d for sulfate

of 7.1 µM) (SHRESTHA *et al.,* 2002), and the Phosphonate-Binding Protein (K_d for methylphosphonate of 1.3 µM (Rizk *et al.,* 2006). However, orthologues of PstS, show affinities to phosphate at the nanomolar level, such as *E. coli* PstS (K_d of 70 nM) (RICO-JIMÉNEZ *et al.,* 2016) and *Pseudomonas aeruginosa* PstS (K_d of 7 nM) (SOLSCHEID *et al.,* 2015)). The reason why we did not get the Kd for *X. citri* PstS could be explained by the high affinity of PstS for PO_4^{3-} and its quick saturation, omitting the perception of the signal through TFS assays, which has a limit for detection in the µM range but not nM (YAMMINE *et al.,* 2019).
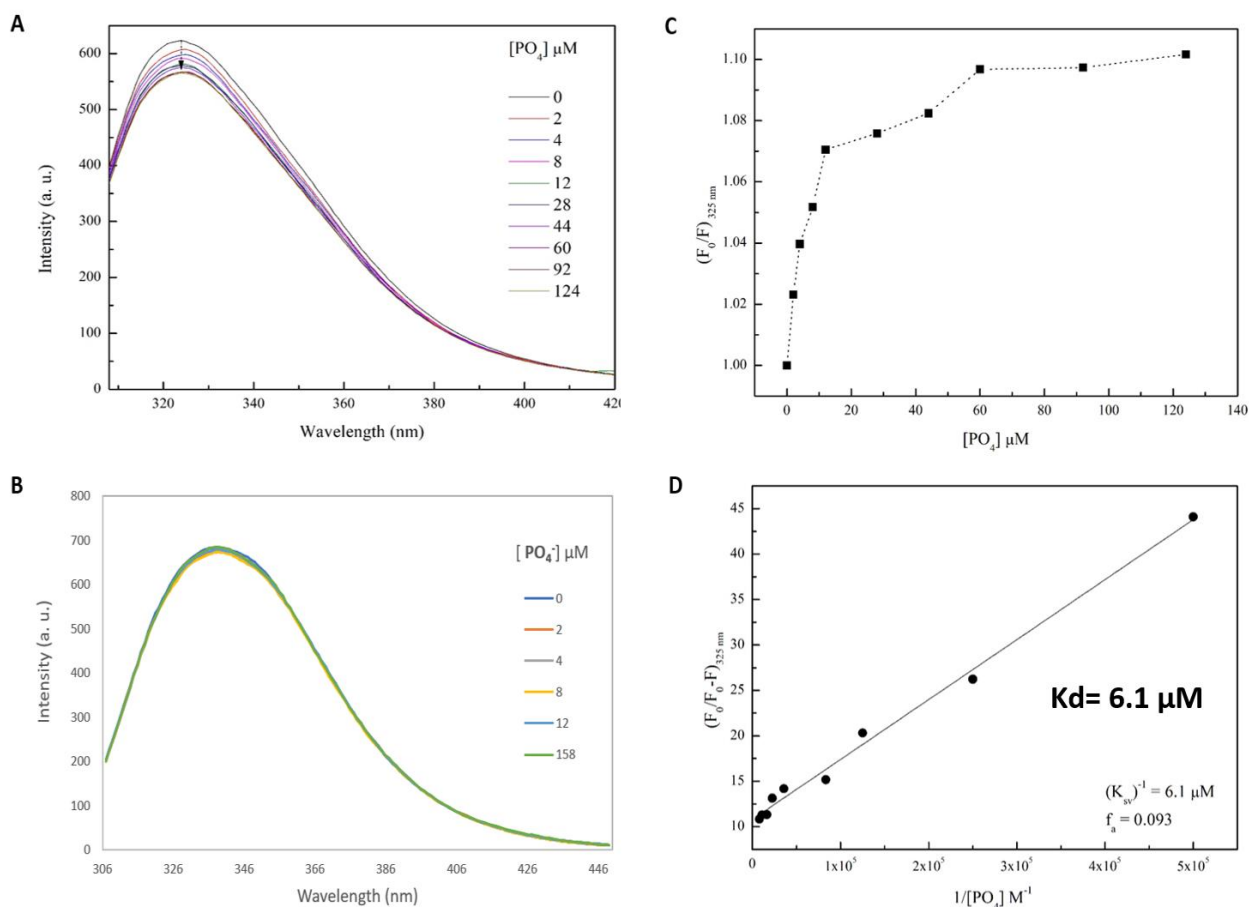
Figure 23. Tryptophan Fluorescence Spectroscopy (TFS) to determinate PstS and PhoX affinities to phosphate. A. Emission spectra of PhoX with increasing concentrations of PO_4^{3-}. It is observed a tryptophan quenching with maximum fluorescence at 325 nm. **B.** Emission spectra of PstS with increasing concentrations of PO_4^{3-}. No quenching is observed. Excitatory light for both proteins at 295 nm. **C.** Stern-Volmer plot showing two populations of fluorophores, one accessible to the suppressor and another with a little accessibility. **D.** The modified Stern-Volmer plot where Kd calculated is 6.1 µM.

4.1.3 Bioinformatic analyses

4.1.3.1 PstS model have the expected characteristics

In view of the fact that one of the objectives of this work was to describe the phosphate uptake by PhoX and PstS, and considering that the biophysical techniques applied to PstS did not give enough information, we decided to carry out molecular dynamic simulations of both proteins in the absence and presence of the anion. For this, since the tridimensional structure of PstS has not been resolved so far, a tridimensional model was generated using the Modeller program, the Clustal Omega tool, and the structural coordinates of *Yersinia pestis* PstS (PDB identification code 2z22) as template. PstS from *X. citri* and *Yersinia pestis* share an amino acid sequence identity of 52.26%, and conserve six of the seven residues that compose the phosphate binding site, including the Asp residue that confers high selectivity respect other anions and the Arg residue that stabilizes the pocket by ionic bonds with other protein residues.

The selected PstS protein model from *X. citri* had a RMSD of 0.091 and a DOPE score of -37682.68 with *Yersinia pestis* PstS, and no residues in regions generously allowed or disallowed in the Ramachandran plot **(Appendix 4)**. So, the tridimensional structure of the model considered 315 amino acids from the 363 that belong to the whole PstS protein (from G49 to K363), since the first 48 amino acids were non-structured. The PstS model shows two domains (Domain I and II) and the hinge region with a loop-helix-loop **(Figure 24A)**, as expected in PBPs (LEDVINA *et al.*, 1998; DWYER *et al.,* 2004; ELIAS *et al.*, 2012; GONZALEZ *et al.*, 2014). In the model, Domain I (or N- Domain) has six helices and five strands, Domain II (or C- Domain) seven helices and six strands, and the hinge region contains the α-helix flanked by loops. Similarly to what was observed in PhoX, PstS model has in domain II a helix surrounded by loops **(Figure 24B**, pink**),** which could experiences movements upon the phosphate bind as suggested in PhoX SAXS results**.**

Comparing the generated PstS model with the available structure of PhoX (PDB:5I84), we observed that the 70% of amino acid sequence identity is reflected in the structural similarities that include the two alpha-beta domains, the hinge region and the

ligand-binding site **(Figure 25A).** On the other hand, the main differences can be related to the number of structures forming helices and sheets that are higher in PhoX. **(Figure 25B)**. The binding pocket is formed by seven amino acids (S58, S86, D104, R183, G188, T189 and N225 in PstS and S34, S62, D80, R159, G164, T165 and N201 in PhoX) (PEGOS *et al.,* 2017), including Asp in Domain I and Arg in Domain II, two residues that will be further mentioned **(Figure 26)**.

Structural model of PstS from *X. citri*
Yersinia pestis PstS (PDB: 2z22)
Identity: 52.26%
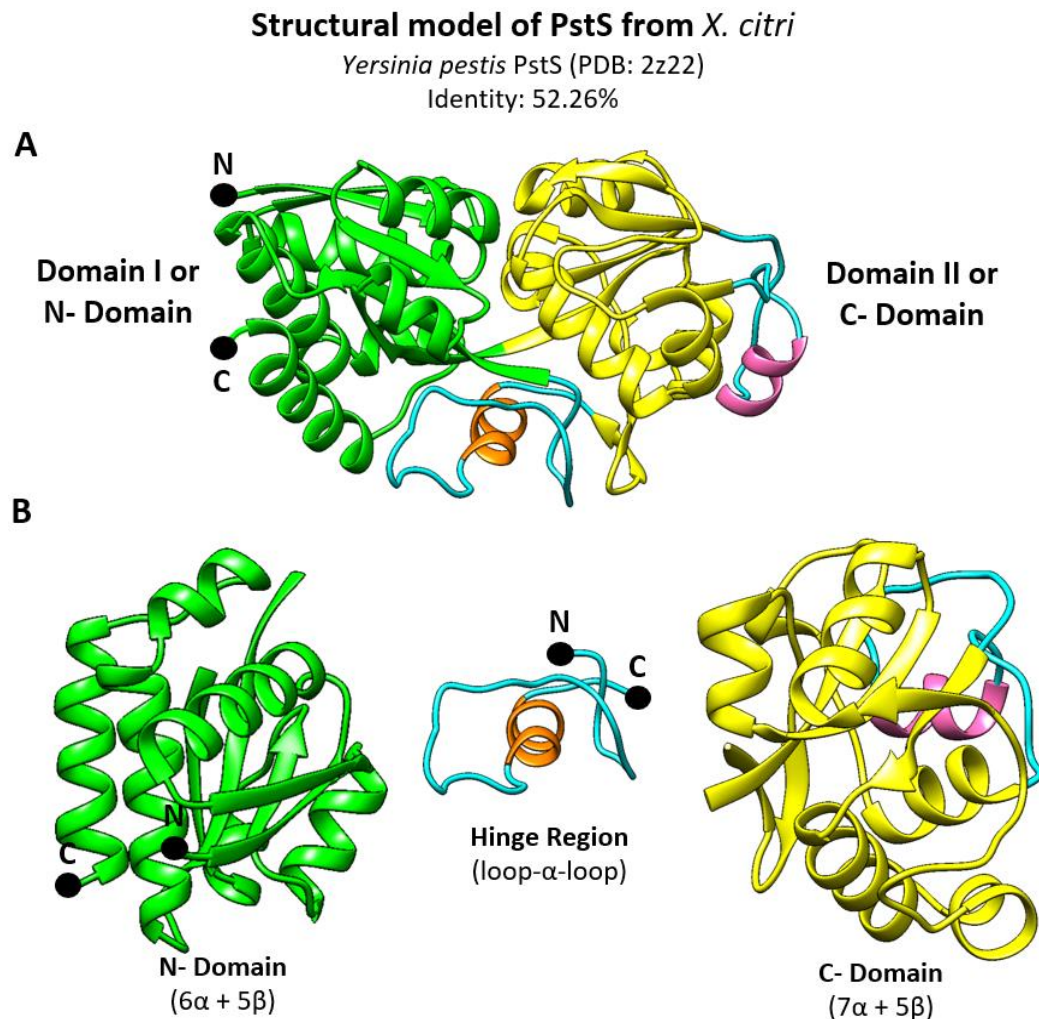
Figure 24. Structure model of *X. citri* **PstS. A.** PstS model has the characteristic two domains (Domain I in green and Domain II in yellow) joined with a hinge region (loop-helix-loop in cyan and orange). **B.** Domain I (green cartoon) has six helices and five strands; Domain II (yellow cartoon) has seven helices and six strands, also a flexible helix connected by loops.
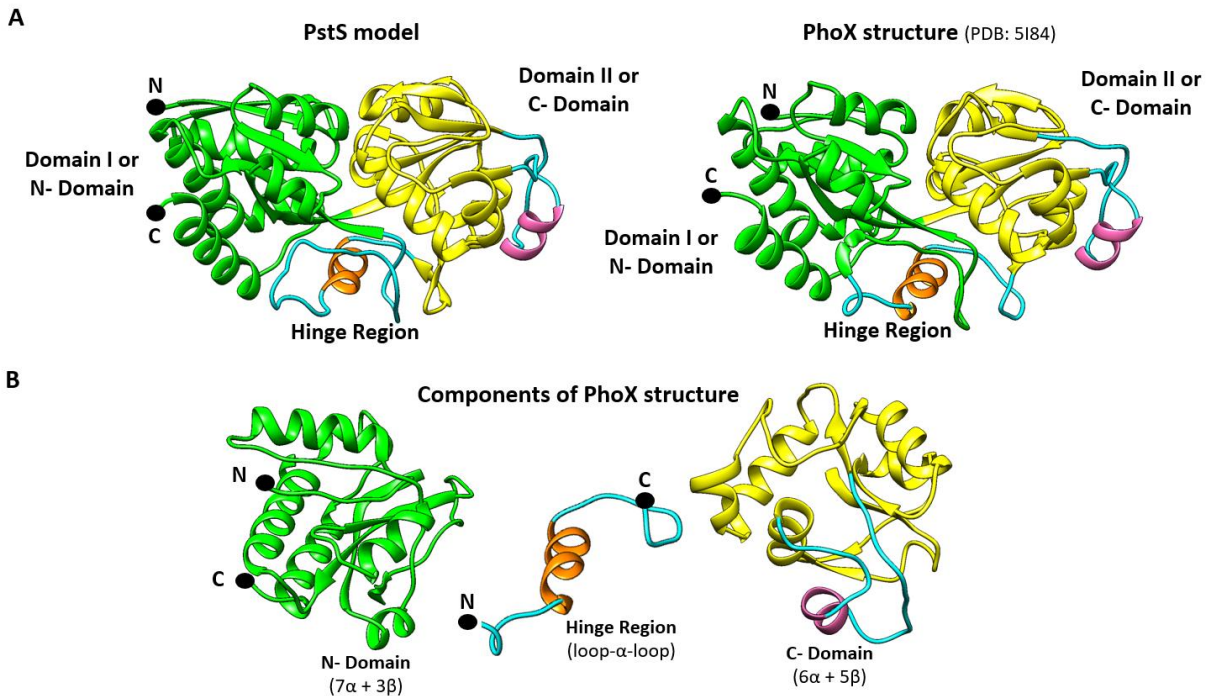
Figure 25. Comparison of PstS model with PhoX structure. **A.** PstS structural model and PhoX crystallographic structure have the same three components: Domain I, Domain II and the hinge region, also in both structures we recognize a helix surrounded by loop in C- Domain that can help in the fit of the ligand. **B.** PhoX components: Domain I (green cartoon) has seven helices and three strands, Domain II (yellow cartoon) has six helices and five strands, and the hinge region (cyan and orange) has a helix connected by loops.
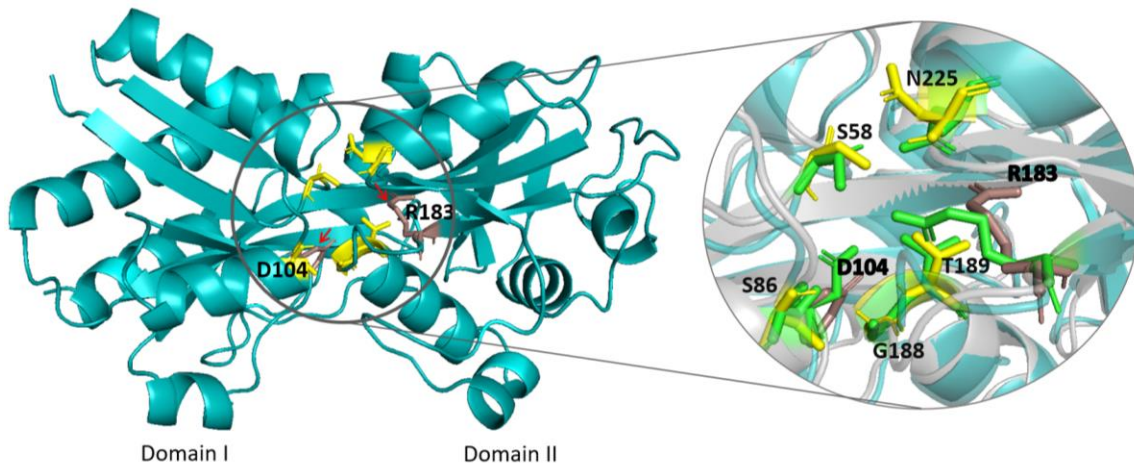
Figure 26. The predicted binding site of *X. citri* **PstS model**. On the left, Pst model depicting as yellow sticks the residues from the phosphate binding site, and as dirty violet sticks the residues D104 and R183 responsible for the selectivity and the formation of an ionic bound with an aspartate in Domain II, respectively. On the right, the overlapping of *X. citri* PstS (in cyan color) and PhoX (in gray color) binding sites, confirming the conservation of the seven residues. In green sticks the residues corresponding to PhoX protein.

4.1.3.2 Molecular dynamic simulations for PhoX and PstS characterization in the phosphate bind and release mechanisms

To evaluate the predict movements of both proteins during phosphate uptake (bind and release), we performed molecular dynamics analyses of PhoX and PstS in presence or absence of phosphate monitored during 1 or 1.5 µs. After the simulations, six different aspects were evaluated: (i) Trp exposure and movement, (ii) the influence of phosphate in the movement of the proteins as a whole in comparison with the apo state, (iii) identification of the flexible regions, (iv) N- and C- Domains behavior, (v) calculation of the free energy between the apo- and bound- proteins, and (vi) behavior and identification of the important residues for phosphate bind and release.

The molecular dynamic simulations (MD) of PstS and PhoX in presence and absence of phosphate were monitored during the 1.5 µs. The position of the tryptophans during the simulations was correlated to the fluorescence observed in the TFS **(Figure 27A, 27B)**. The MDs revealed movement of W135 in PhoX, which is part of one of the loops around the alpha-helix in C-Domain that is highly flexible **(Figure 27C)**. The quenching of the fluorescence could be induced by movement of this residue to hidden location. In PstS, W159 is the corresponding residue to W135 but accordingly to the structural model, W159 does not have as much freedom for movement since it is surrounded by a beta-sheet that hides it from the surface **(Figure 27D)**. The MD also showed that tryptophans in PhoX are more hidden in the presence of phosphate than absence which in consonant with the TFS results. However, in the case of PstS the tryptophans exposure is fluctuating.
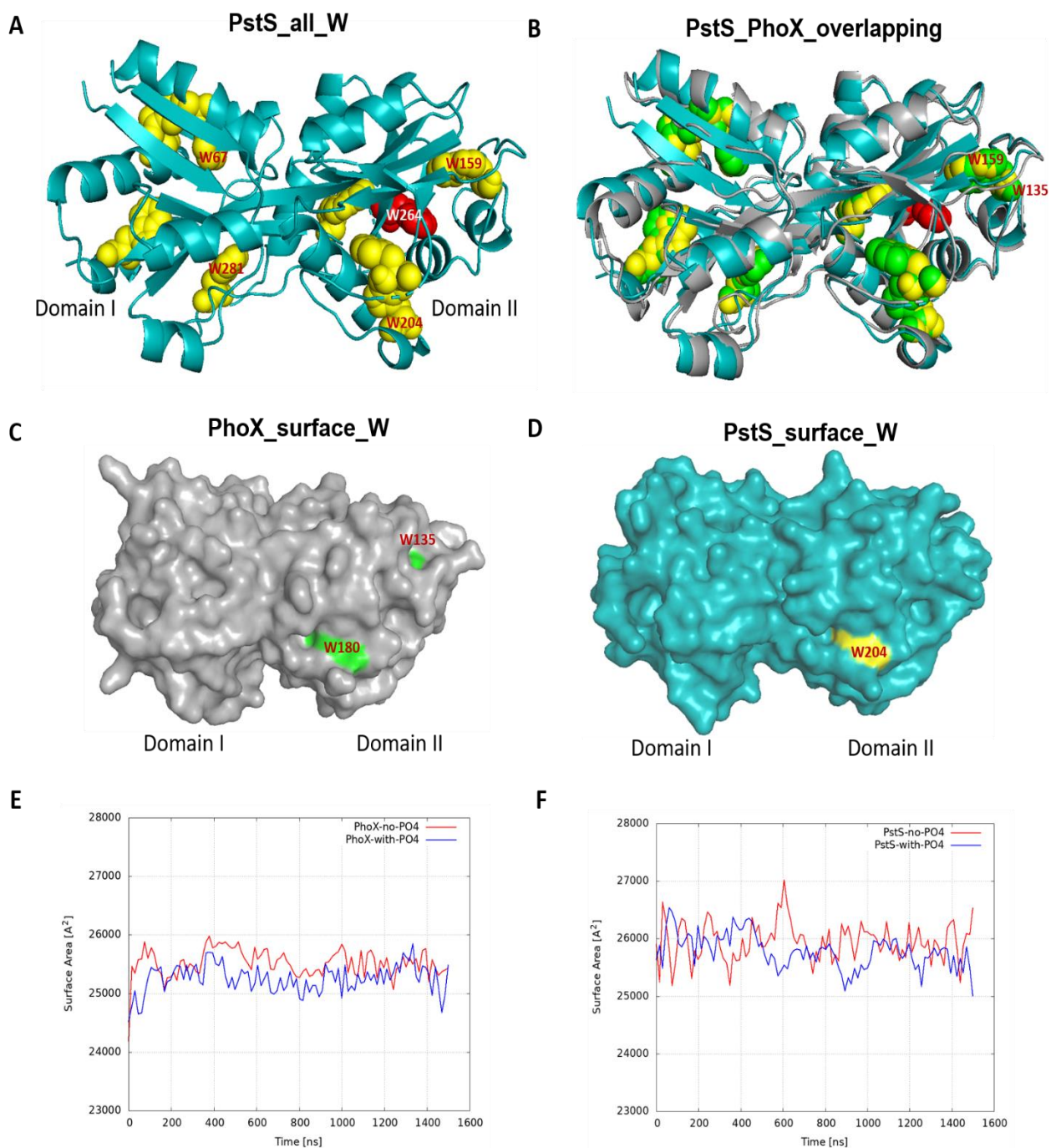
Figure 27. Tryptophan location and exposure in PhoX and PstS during phosphate binding. A. PstS structural model in cartoon representation (cyan color) showing the 8 conserved tryptophan (W) in yellow and the exclusive W264 in red spheres. **B.** Superimposition of PhoX (gray color) structure and PstS model (cyan color) highlighting the conservation of the 8 Trp (spheres). **C.** Three-dimensional structure of PhoX bound to phosphate is shown as surface. The position and exposure of W180 and W135 are evidenced in green. **D.** PstS model represented as surface showing W159 (which is the pair of W135 from PhoX) hidden. **E.** Plot of Trp exposure in PhoX showing a decrease of exposure when PhoX is phosphate bind. **F.** The plot of Trp exposure in PstS shows fluctuation in its apo- and bound- states.

Following the MD analysis, we monitored the behavior of the whole proteins PhoX and PstS. For this purpose, it was evaluated two variables: the angle formed between the three centers of mass from the ligand, C- and N- Domains, and the distance between C- and N-Domains. The results indicate that these variables in PhoX and PstS are maintained in a constant movement, and are more stable in the presence of phosphate. Comparatively, PhoX is more stable than PstS, which had parameters continuously oscillating reflecting the opening and closing of the domains. The addition of PO_4^{3-} stabilize the protein and keeps the parameters stable **(Figure 28A, 28B)**. The calculated RMSD (Root Mean Square Deviation) between PhoX and PstS proteins in the bound and unbound states were 1.6 Å and 1.8 Å, respectively. It corroborates that PhoX is more stable than PstS and have restricted movements of N- and C- Domains **(Figure 28**, superimposed structures**)**.
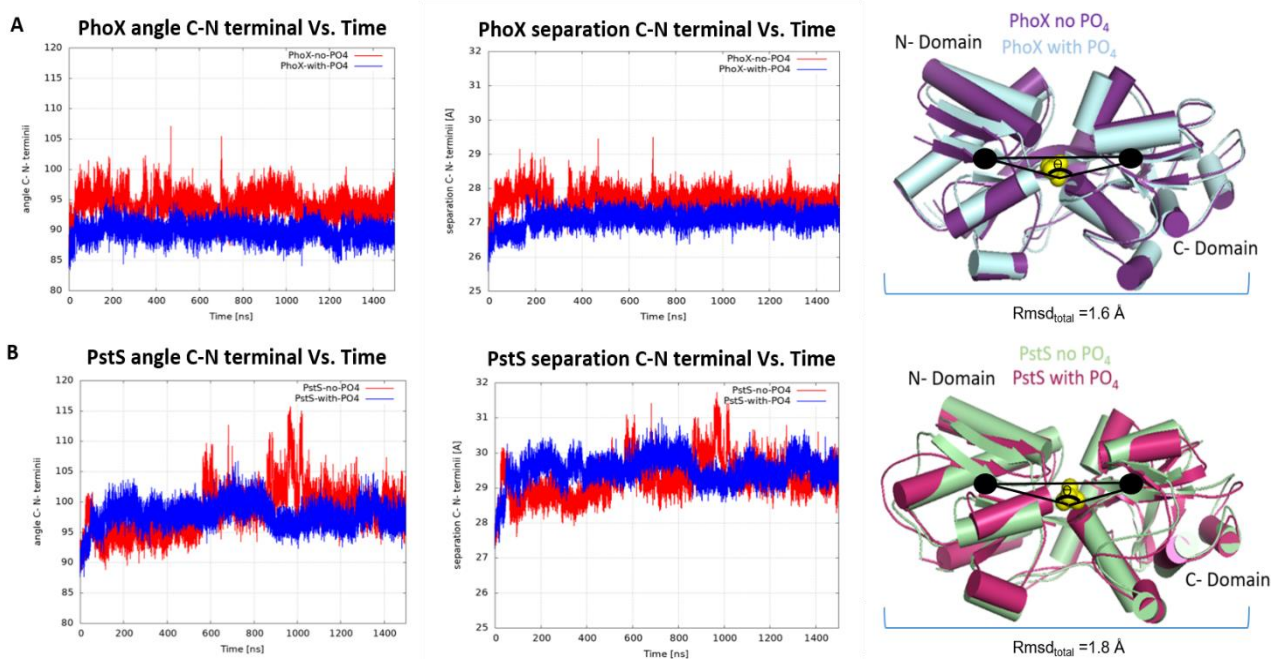
Figure 28. PhoX and PstS behavior during molecular dynamic simulations in the absence and presence of phosphate. A. Profiles of PhoX and PhoX-PO_4^{3-} angles and distances between N and C-termini during the simulations showing a decrease in the angle and distance between domains in presence of phosphate. On the left, superimposition of PhoX and PhoX- PO_4^{3-} structures indicating the RMSD value of 1.6 Å, and the reference points to calculate distances and angles. **B.** Profiles of PstS and PstS- PO_4^{3-} angles and distances between N and C-termini. The plots reveal higher oscillation in PstS in absence of phosphate than in the presence. On the left, superimposition of PstS and PstS- PO_4^{3-} structures indicating the RMSD value of 1.8 Å.

The analysis of the MD simulations allowed the determination of the most flexible regions of the proteins. The results were presented in graphs of Residues vs. RMSF (Root Mean Square Fluctuation).

Corroborating the general behavior of the proteins, residues in PstS showed higher flexibility than in PhoX. Moreover, we observed that residues with higher flexibility tend to move for the ligand uptake and coordination, mainly located in the hinge region and C-Domain **(Figure 29A, 29B)**.

In fact, the residues with higher flexibility in the apo-forms of PstS and PhoX are part of or close to the phosphate binding site (S62 in PhoX and F59 in PstS), and are part of N- Domain. On the other hand, in the bound-forms the flexibility is higher in the hinge region (T259 in PhoX and R285 in PstS), and the C-Domain (G111 and P148 in PhoX, and Q215 in PstS), as expected for SBP proteins (RAVI et al., 2007). In the case of PhoX, P148 is part of the helix flanked by loops present in the C- Domain **(Figure 29C, 29D)**.
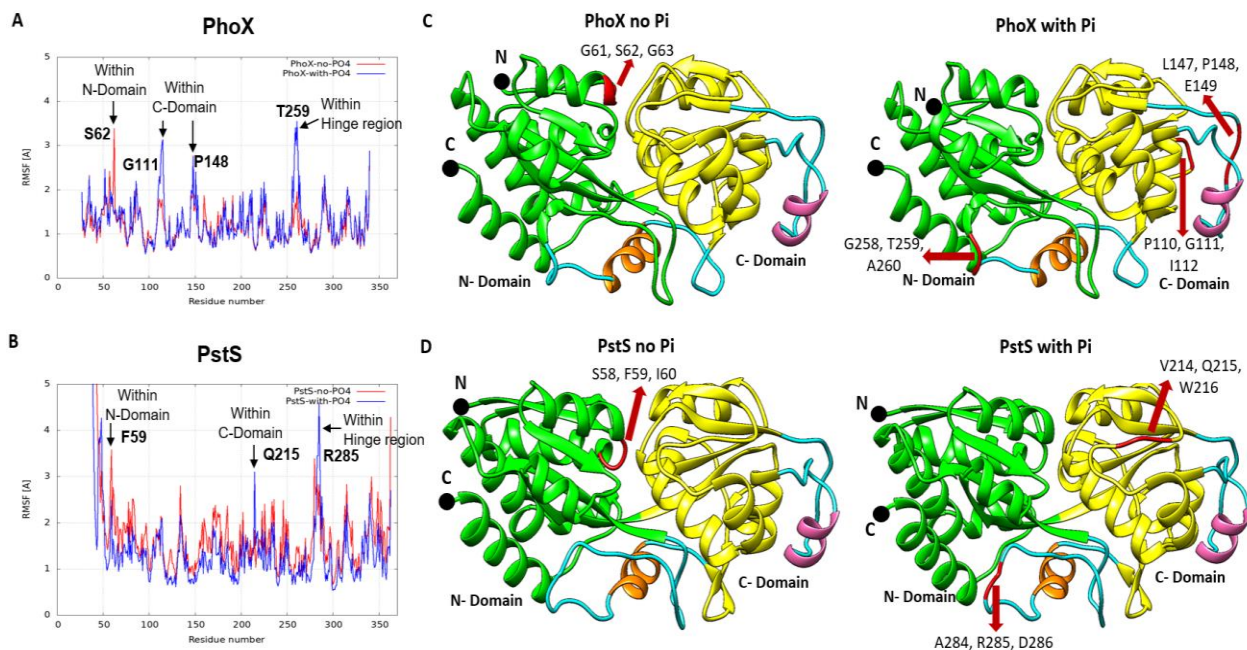
Figure 29. Evaluation of the flexible regions of PhoX and PstS in their apo- and bound- states. A and **B.** Regions and residues from PhoX and PstS with high flexibility in the apo- (in red color) and bound states (in blue color). **C** and **D.** Tridimensional structures of PhoX and PstS proteins without or with phosphate indicating the residues with high flexibility.

We calculated the RMSD (Root Mean Square Deviation) of the N- and C- Domains (separately) comparing the initial and final structures of PhoX and PstS in presence or absence of phosphate. The results indicates that both proteins loss flexibility in N-Domain in presence of phosphate in comparison to C- Domain (see RMSD of Ni/Nf and Ci/Cf, **Figure 30A, 30B**) but clearly the phosphate induces domains closing and limits the movements of residues. It because the total RMSD decrease from 2.7 Å to 2.3 Å in PhoX, and from 2.2 Å to 1.9 Å in PstS (see total RMSD, **Figure 30A, 30B**). These results are consonant with the previously obtained for RBP where the ligand binding enabled the binding pocket turned into more stable form (LI *et al.,* 2021).
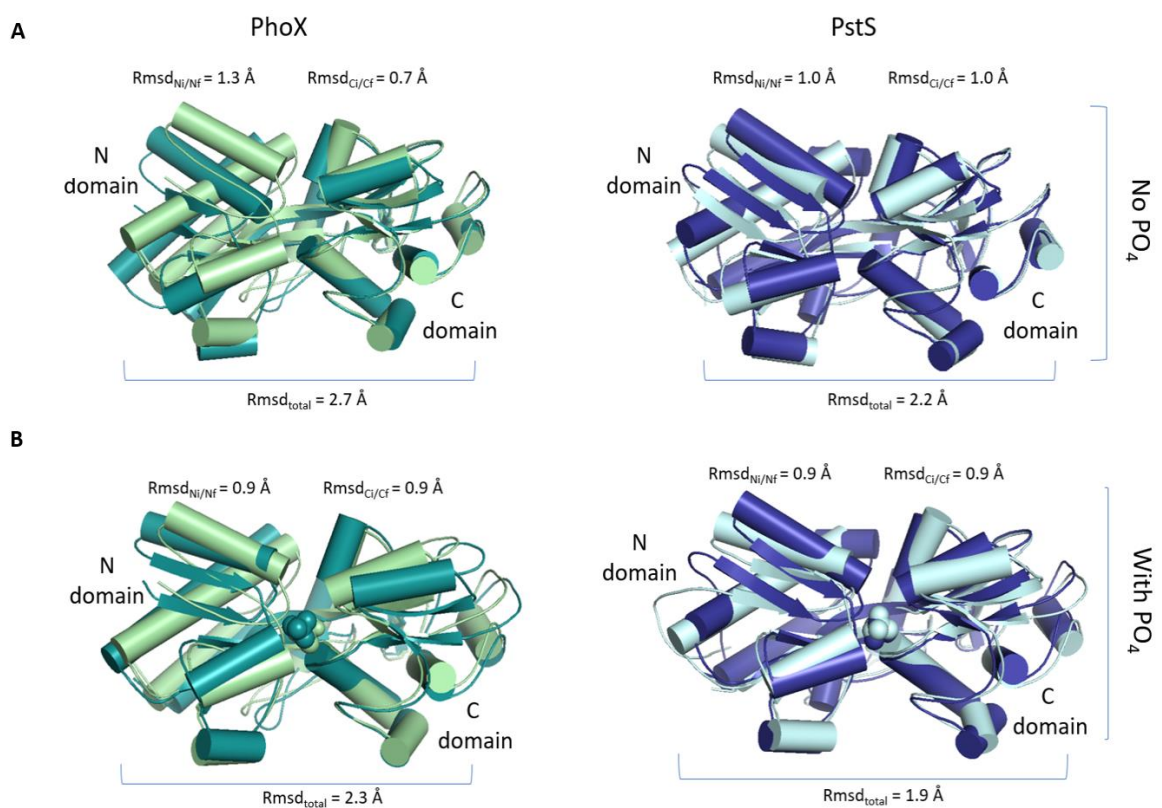
Figure 30. Behavior of N- and C- Domains during the molecular dynamic simulations of PhoX and PstS in absence and presence of phosphate. A. Comparison of RMSD of PhoX and PstS without PO_4^{3-} in the beginning and end of the dynamics in the N- Domain ($Rmsd_{Ni/Nf}$), C- Domain ($Rsmd_{Ci/Cf}$) and the whole protein ($Rmsd_{total}$). **A.** Comparison of RMSD of PhoX and PstS with PO_4^{3-} in the beginning and end of the dynamics of the domains separately and the whole protein.

The free energy differences (DeltaF) of PhoX and PstS interacting with phosphate were calculated. This parameter could be related to the affinity of both proteins for phosphate. DeltaF for both proteins were calculated in -8.0 kcal/mol for $PstS_{DeltaF}$ and ~7.1 kcal/mol for $PhoX_{DeltaF}$. Based on these results PstS seems to have slightly higher affinity to phosphate than PhoX, with an approximately Kd of ~1.6 µM and ~7.3 µM, respectively. According to the literature, PBPs belonging to the ABC transporters have an expected Kd in the submicromolar range, nevertheless until now there is no experimental data about the Kd of *X. citri* PstS and the obtained Kd for *X. citri* PhoX is the first experimental data for this protein.

Finally, we performed an evaluation of the residues involved in the phosphate binding site of PhoX and PstS during the simulations. The analysis showed that 9 residues are involved in the phosphate interaction of PhoX (S34, S62, D80, R159, S163, G164, T165, N201, Y222) and PstS (S58, S86, D104, R183, G186, S187, G188 T189, N225). From these, S163 and Y222 from PhoX, and G186 and S187 from PstS are additional to the 7 conserved residues reported for both protein by PEGOS et al. (2017). All these residues are distributed in four helices, two in each domain, which constitute the local dipoles surrounding the buried ligand. Moreover, we identified the aspartate that form the ionic bond with a specific arginine (from the binding site) in the C-Domain. These residues in PhoX and PstS were: D161 with R159 and D185 with R183, respectively. The MD simulation of PstS and PhoX in the beginning and end of the dynamics revealed the importance of residues belonging to the dipole that coordinates the anion. These residues tend to push the ligand to the entrance of the pocket, S62, D80, R159 and T165 in PhoX and S86, R183, G188 and T189 in PstS **(Figure 31A, 31B).** In the entrance of the pocket, residues R159 (in PhoX) or R183 (in PstS) possibly play an important function for ligand delivery.
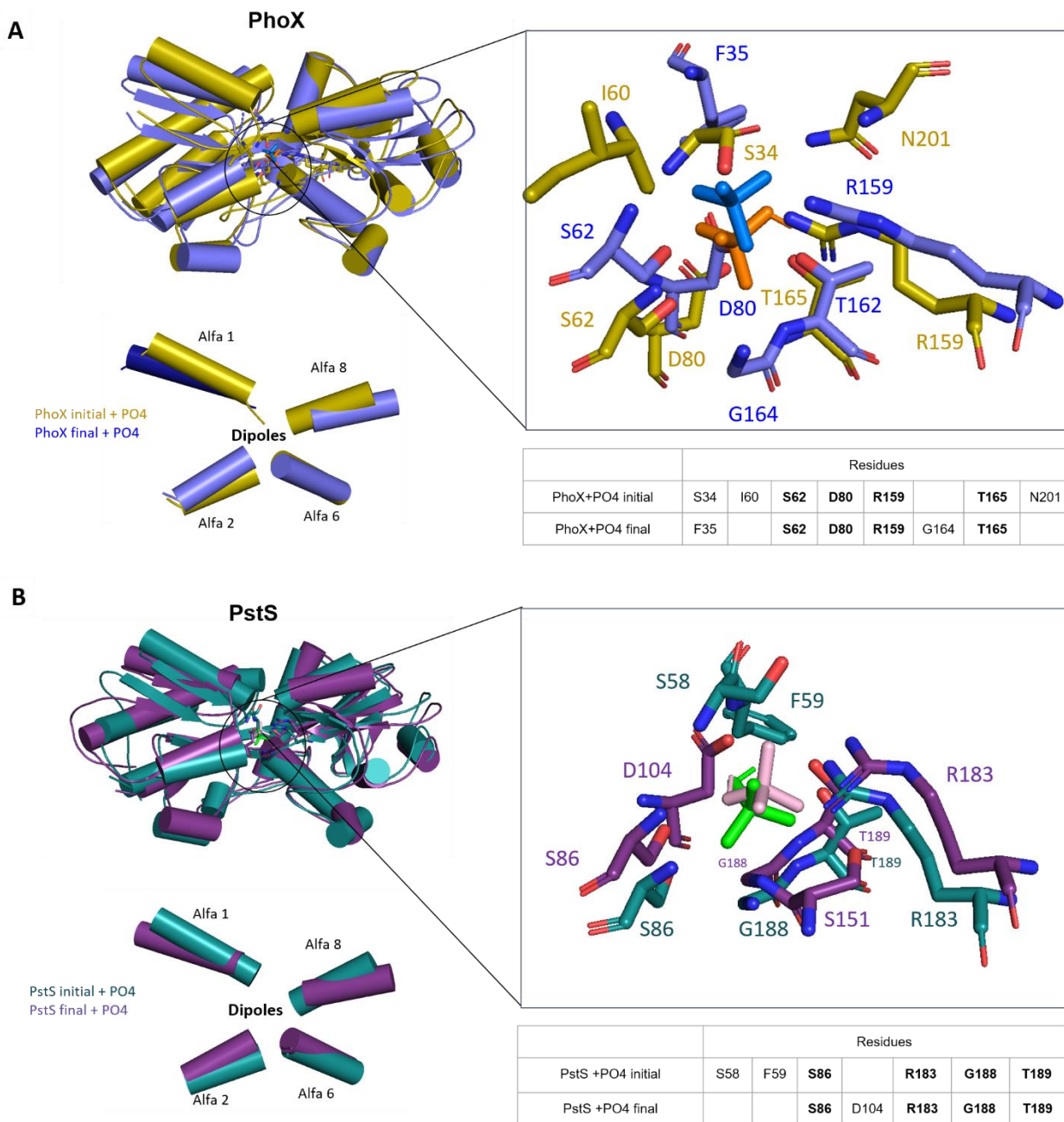
Figure 31. Trajectory of the residues forming the ligand-binding site of PhoX and PstS from the beginning to end of the dynamics in presence of phosphate. A. PhoX characteristics. The initial (yellow) and final (blue) structures of PhoX are superimposed in cartoon mode and the residues from the ligand-binding pocket are shown in detail (sticks). The phosphate anion is also shown. In the structures, is possible to identify 4 helices, 2 forming a dipole (alpha1 and 8) for anion coordination and 2 supporting the residues in the interaction. A comparison among residues that coordinate the anion at the beginning and end of the MD is shown in the table inset. Residues S62, R80, G159 and T165 are responsible for pushing the phosphate to the entrance of the pocket. **B.** PstS characteristics. Similarly, PstS is shown in cartoon representation, with residues in detail (sticks). The helices that coordinate the phosphate are shown as well as the comparison of the residues. Residues S86, R183, G188 and T189 are responsible for pushing the phosphate to the entrance of the pocket.

Based on the MD simulations we propose a model for PstS and PhoX uptake and phosphate delivery mechanisms. In this model, N- Domain would play for first ligand interaction, with play of a serine and an aspartic acid that also might be responsible for the selectivity and the anion attraction. The flexibility decreases in the serine or close residues in N- Domain, and the flexibility increases in the hinge region and the C-Domain helps the fitting of the phosphate into the pocket, which is supported by an interaction with an arginine from the C- Domain and an ionic bound between this arginine and an aspartate in the same domain. The movement of these charged residues would induce the ligand shift to the entrance of the pocket, needed for its delivery **(Figure 32).** Interestingly, these residues are highly conserved in other anion-binding proteins.
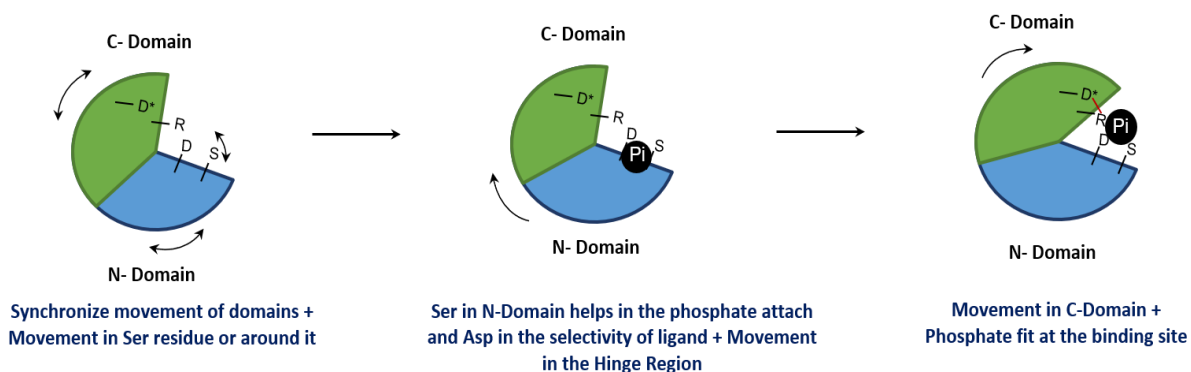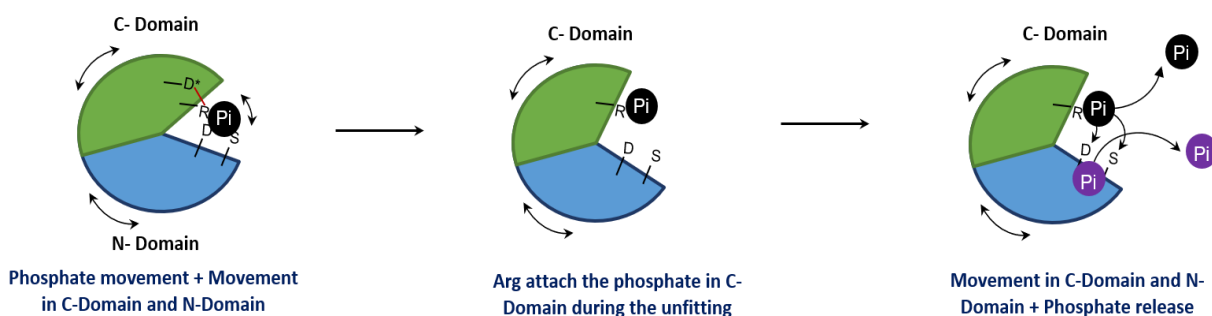
Figure 32. Model of PBPs behavior during the phosphate uptake (bind and release). Charged residues are important for attraction of the ion for binding and delivery.

In the deposited three-dimensional structure of *X. citri* PhoX (PDB 5I84) there are three phosphates bound. Two of them are positioned in the interface in the way to the pocket and the third one inside the pocket. That fitting suggested us that this could be the route for phosphate. To evaluate that possibility, we performed MD simulations looking at the phosphate moving to the ligand-binding pocket. For this purpose, the second PO_4^{3-} closer to the pocket was chosen to start the simulation. However, that phosphate moved far away from the pocket to K70 and K199 and the it was released. The presence of lysines in the entrance of the pocket and the phosphate attraction called our attention, especially because these two lysines are conserved in PstS with K94 and K223, respectively **(Figure 33A).** These residues are located in the interface of the periplasmic-binding proteins that interact with the permeases and confer a positive charge.

Complementing the analysis, using the AlphaFold server, we built a structural model of *X. citri* permeases PstA and PstC based on the three-dimensional permeases from the *Archaeoglobus fulgidus* molybdate transporter (HOLLENSTEIN *et al.,* 2007; PDB 2ONK). Interestingly, the surface electrostatic potential of PstA and PstC in the region that interacts with the PBPs, revealed negatively charged spots, which might fit the mentioned lysine. Similarly, the molybdate-binding protein ModA also conserves positive residues in similar positions, R82 (in the N- Domain) and R215 (in the C-Domain), which have complementary charges in the permeases ModB **(Figure 33B).**

We also performed a structural alignment of PstS and PhoX (PDB: 5I84) with other phosphate-binding proteins that had solved three-dimensional structures: *Mycobacterium tuberculosis* PstS1 and PstS3 (1PC3 and 4LVQ, respectively), *Clostridium perfringens* PBP (4Q8R), *Pseudomonas fluorescens* PBP (4F1V), *Yersinia pestis* PstS (2Z22), *E. coli* PstS (1IXH) and *Vibrio cholerae* SBP (1TWY). All of them showed the conservation of the two positively charged residues (arginine or lysine) in the region that interacts with the permeases **(Figure 33C).** These results suggest that the positive residues in PBPs play important role in the interaction with the permeases.

In this manner, we stablish our first hypothesis, these lysines help in the charge neutralization of the permeases allowing the entrance of phosphate from the PBPs to the inner of the permeases (as part of the phosphate uptake) by charge affinity between the ligand and the pocket in the permeases.
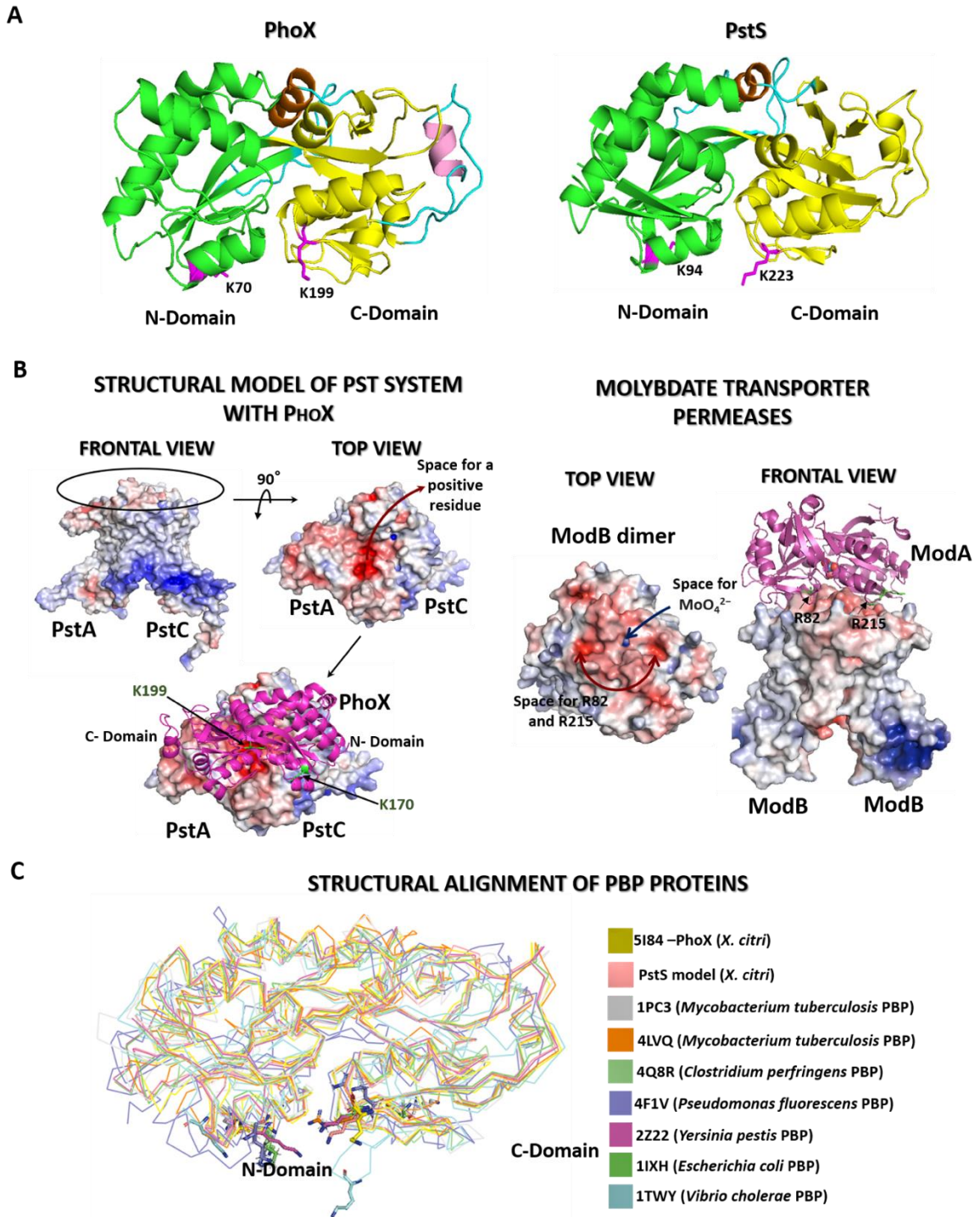
Figure 33. Hypothesis of the role of specific lysines in the phosphate transport. A. Location of lysine in PhoX and PstS that bind phosphate for few nanoseconds before its release. **B.** Comparing the frontal and top view of PstA and PstC permeases (from Pst system) with the ModB permeases (from the molybdate transporter), both systems have a clearly negatively region in the surface where putatively, the positive lysine or arginine could interact. **C.** A structural comparison of nine PBPs showing the structural conservation of positive residues in the region of interaction with permeases.

A second hypothesis raised from these analyses would be that mentioned residues in PBPs could work as sensors of the phosphate levels in the periplasm. In high concentration of PO_4^{3-}, all the PBPs would have PO_4^{3-} at the binding site and bound to the lysines (based on the MD results), impairing the fitting in the permeases. In low concentration of the anion, the high affinity of the PBPs would keep PO_4^{3-} only at the binding site, and with the positive residues are free to fit in the permeases **(Figure 34)**. This hypothesis also suggests that PBPs are not permanently attached to PstA and PstC **(Figure 35)**.

Certainly, these hypotheses need to be experimentally confirmed.
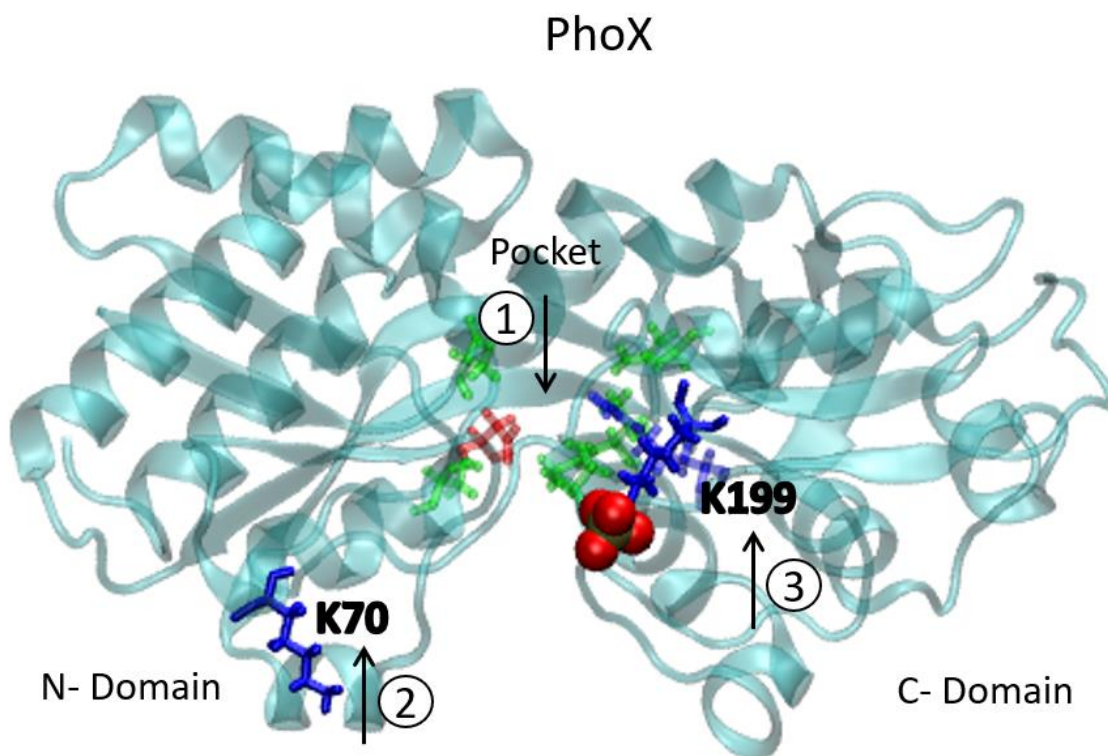
Figure 34. Possible positions for phosphate attaching in PhoX. Three-dimensional structure of PhoX in cartoon showing the three positions where PO_4^{3-} ion can bound: The pocket (in green sticks) and two lysines, K70 and K199 (in blue sticks) located in the N-Domain and C-Domain, respectively. According to our hypothesis, in high concentration of PO_4^{3-} the three positions will be phosphate bound, while in low concentration of PO_4^{3-}, only the binding site would have phosphate bound due to the high affinity.
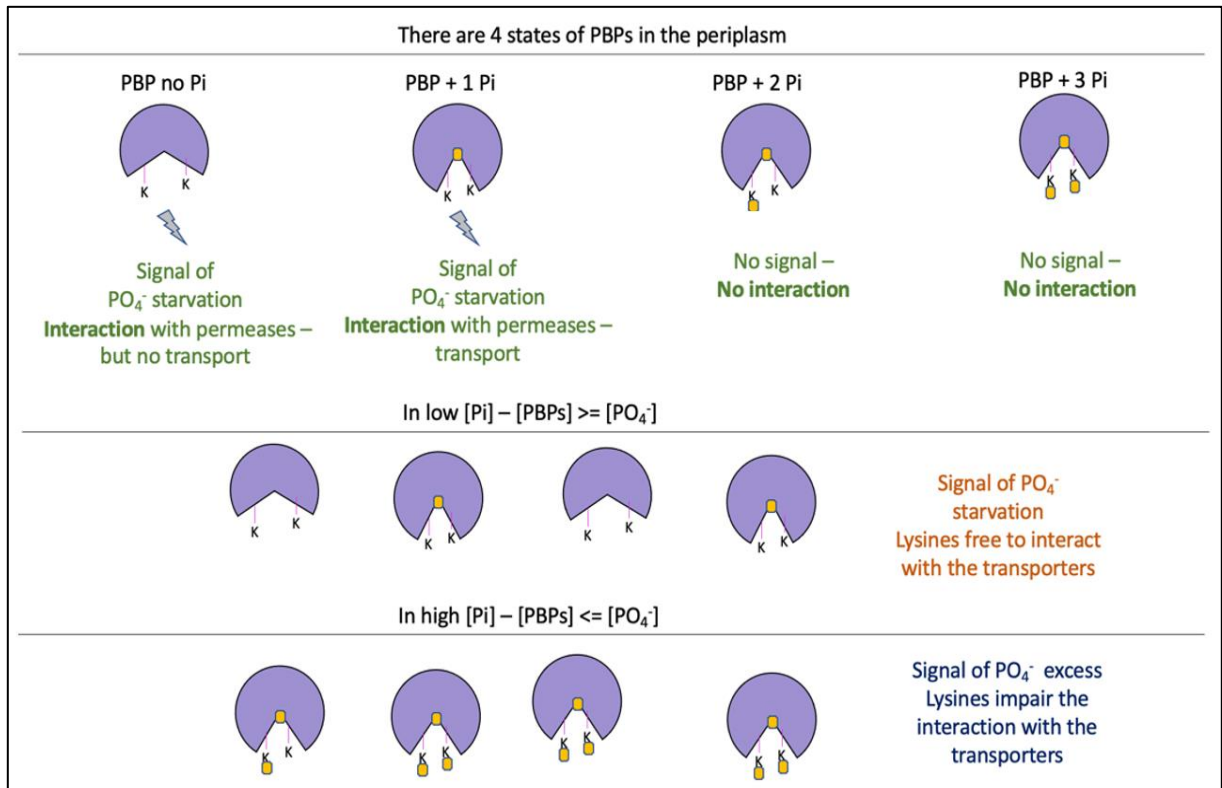
Figure 35. Hypothesis of the sensing of phosphate level and transport activation mediate by the lysine.

4.1.4 Conditions for obtaining PstS crystals

Since there is no available structure from *X. citri* PstS, we worked on the crystallization of PstS. We used two concentrations of proteins, 10 mg.ml^{-1} and 23.5 mg.ml^{-1} in the hanging-drop vapour-diffusion and the sitting-drop vapour-diffusion methods, respectively. In the assays the molar ratio of protein: buffer was 1:1. The plates were incubated at 18°C. PstS crystals were obtained with 23.5 mg.ml-1, in different conditions. All crystals showed thin needle shape that need to be refined **(Figure 30).** As a conclusion, we can say that the best precipitant for PstS is Ammonium Sulfate in concentrations between 1.4 to 2.4 M.
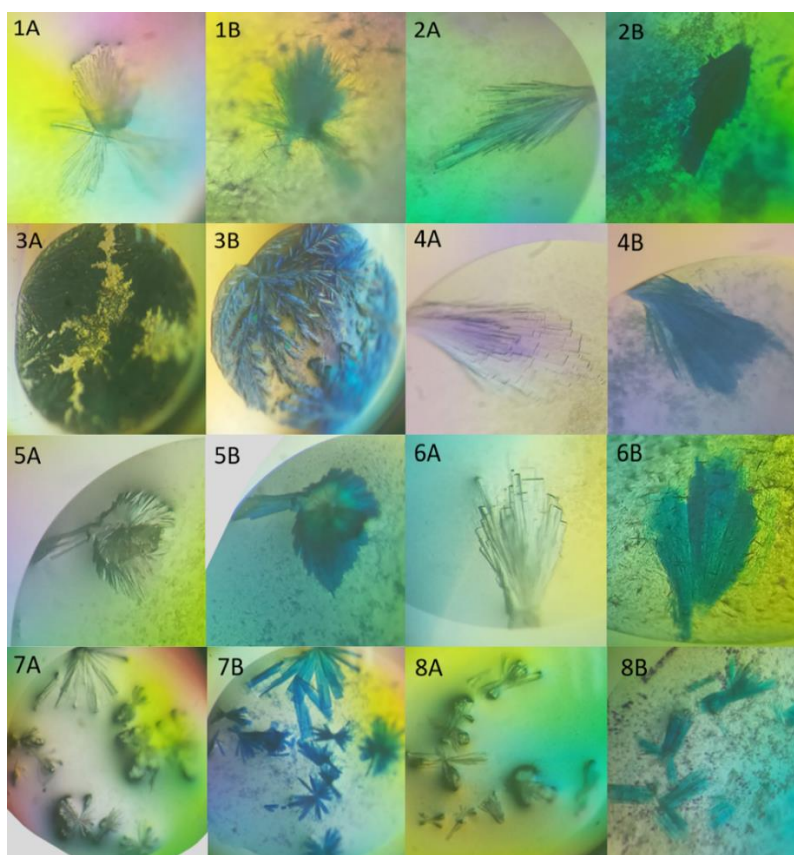
Figure 36. PstS crystals. A. PstS crystals reached in eight different buffers: 1. 150 mM Tris pH 8.0 buffer, 1.60 M Ammonium Sulfate and 28 mM Sodium Fluoride (1A), 2. 150 mM Formate pH 4.0 buffer, 2.40 M Ammonium Sulfate and 3.0% v/v 2-Butanol (2A), 3. 150 mM AMPD-Tris pH 9.0 buffer, 1.40 M Ammonium Sulfate and 8.0 % v/v MPD (3A), 4. 5% w/v 2-Propanol and 2 M Ammonium Sulfate (4A), 5. 100 mM MES pH 7.5 buffer, 1,6 M Ammonium Sulfate and 10% v/v 1,4-Dioxane (5A), 6. 100 mM Tris pH 8.5 buffer and 2 M Ammonium Sulfate (6A), 7. 1.6 mM TRI-Sodium citrate pH 6.5 buffer (7A) and 8. 2 M Ammonium Sulfate (8A). **B.** PstS crystals in the eight different buffers mentioned in A and colored by a solution of 0.05% (w/v) of methylene blue.

4.2 Chapter II

4.2.1 Bioinformatic analyses of chemotactic proteins

As mentioned in the introduction, PstS in *P. aeruginosa* plays an additional role in the chemotaxis, interacting with the ligand-binding domain (LBD) of two chemotactic proteins: CtpH (PA2561) and CtpL (PA4844). To search the putative orthologues of these proteinsin *X. citri*, a basic local alignment with Blastp was performed using their amino acid sequences against *X. citri* databank. The search resulted six proteins putatively involved in chemotaxis: Mcp1 (XAC3213), Tsr1 (XAC1896), Tsr2 (XAC0611), Tsr3 (XAC1897), Tsr4 (XAC1892) and Tsr5 (XAC1891) **(Table 1)**. To proceed with the studies, we selected only Tsr1 and Tsr2 since they were properly mentioned in the genome annotation as chemotactic proteins of the MCP family, besides to have the highest amino acid sequence coverage and identity compared to the *P. aeruginosa* proteins.

Table 1. Orthologs of *P. aeruginosa* **CtpH and CtpL proteins in** *X. citri.* CtpH orthologs (at the top of the table) and CtpL orthologs (at the bottom of the table). The table shows the coverage, identity and function of each ortholog (UniProt information).

Blastp of *P.aeruginosa* CtpH against *X. citri*				
Microorganism	Protein	Coverage	Identity	Function
P. aeruginosa	CtpH (PA2561)	100%	100%	Chemoreceptor for inorganic phosphate, which is required for taxis at high concentrations of phosphate. It recognizes inorganic phosphate directly and can bind other components that have a pyrophosphate group, including ATP and ADP.
X. citri	Mcp1 (XAC3213)	58%	33%	Signal transducer activity and chemotaxis.
	Tsr1 (XAC1896)	71%	31%	Signal transducer activity and chemotaxis.
	PilJ (XAC3099)	81%	31%	Signal transducer activity and chemotaxis.
	Tsr2 (XAC0611)	55%	33%	Signal transducer activity and chemotaxis.
	Tsr3 (XAC1897)	59%	32%	Signal transducer activity and chemotaxis.
Blastp of *P.aeruginosa* CtpL against *X. citri*				
Microorganism	Protein	Coverage	Identity	Function
P. aeruginosa	CtpL (PA4844)	100%	100%	Chemoreceptor for inorganic phosphate, which is required for taxis at low concentrations of phosphate. Does not recognize inorganic phosphate directly, but via a complex between the periplasmic protein PstS and inorganic phosphate.

	PilJ (XAC3099)	62%	32%	Signal transducer activity and chemotaxis.
	Tsr2 (XAC0611)	83%	33%	Signal transducer activity and chemotaxis.
X. citri	Tsr4 (XAC1892)	70%	29%	Signal transducer activity
	Tsr5 (XAC1891)	63%	27%	Signal transducer activity
	Tsr1 (XAC1896)	61%	29%	Signal transducer activity and chemotaxis.

Predictions in TOPCONS server revealed that *X. citri* Tsr1 and Tsr2 proteins show similar secondary structures to CtpH and CtpL proteins. The prediction showed an extracellular domain and an intracellular domain connected by a transmembrane helix. Using the DAS–Transmembrane Prediction server, the Ligand Binding Domain (LBD) from Tsr1 (Tsr1-LBD) and Tsr2 (Tsr2-LBD) were located in the outside region of the extracellular domains. The amino acid sequence alignment of Tsr1-LBD and Tsr2-LBD with *P. aeruginosa* CtpH-LBD and CtpL-LBD evidenced that they share low sequence identity (12 to 17%), which is in accordance with the sequence identity evidenced among proteins located in the periplasm **(Table 2)**.

Table 2. Comparison of the **LBD (Ligand Bind Domain) of** *P. aeruginosa* **CtpH and CtpL and** *X. citri* **Tsr1 and Tsr2.** The sequential identity between the proteins CtpH-LBD and CtpL-LBD are low, in the range of 12 to 17%.

	X. citri proteins (% identity)			
P. aeruginosa	CtpH-LBD	CtpL-LBD	Tsr1-LBD	Tsr2-LBD
CtpH-LBD	100%	17%	14%	12%
CtpL-LBD	17%	100%	15%	13%

4.2.2 Chemotactic proteins Tsr1 and Tsr2 are partly soluble

Fragment of the genes *tsr*1 and *tsr*2 encoding the Tsr1 and Tsr2 LBDs (Tsr1-LBD and Tsr2-LBD proteins) were successfully cloned in pOP3BP plasmids. After the restriction assays of pOP3BP-Tsr1$_{LBD}$ and pOP3BP-Tsr2$_{LBD}$ plasmids with *Bam*HI and *Hin*dIII enzymes, the gene fragments of 378 pb (*tsr*1) and 384 pb (*tsr*2) were released and visualized in an agarose gel **(Figure 37)**.

Tsr1-LBD and Tsr2-LBD proteins showed a molecular mass of 25.77 KDa and 25.92 KDa, respectively. After the overexpression of both proteins in *E. coli* BL21(DE3) cells,

Tsr1-LBD was mostly expressed in the soluble fraction, and Tsr2-LBD was found in significant amounts in the insoluble fraction (pellet) **(Figure 38A, 38B)**.

The fractions obtained from Immobilized Nickel-Affinity Chromatography (IMAC) were analysed in SDS-PAGE and eluted with imidazole concentrations between 100 and 148 mM **(Figure 38C, 38D)**.

To reduce the contaminants and remove the imidazole, both proteins were purified by Size Exclusion Chromatography (SEC) **(Figure 38E, 38F)**. After cleavage of GB1 fusion protein from Tsr1-LBD and Tsr2-LBD, the expected molecular masses were 14.387 KDa and 14.532 KDa, respectively **(Figure 38G, 38H)**. At the end, the recovery of each protein was 0.5 mg per 4 liters of culture medium. Due to the long process to obtain the proteins and the low production efficiencies, we decided to stop the work with these chemotactic proteins.
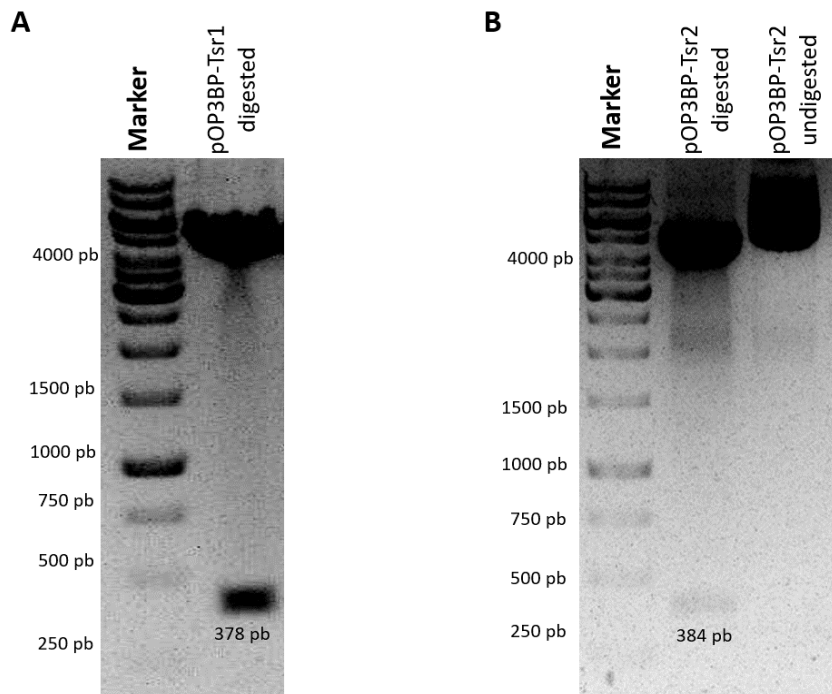
Figure 37. Agarose gel showing pOP3B-Tsr1 and pOP3B-Tsr2 plasmids after restriction assay. **A.** *tsr*1 gene fragment of 378 pb released after the digestion of pOP3B-Tsr1 plasmid with *Bam*HI and *Hind*III enzymes **B.** *tsr*2 gene fragment of 384 pb released after the digestion of pOP3B-Tsr2 plasmid with *Bam*HI and *Hind*III enzymes.
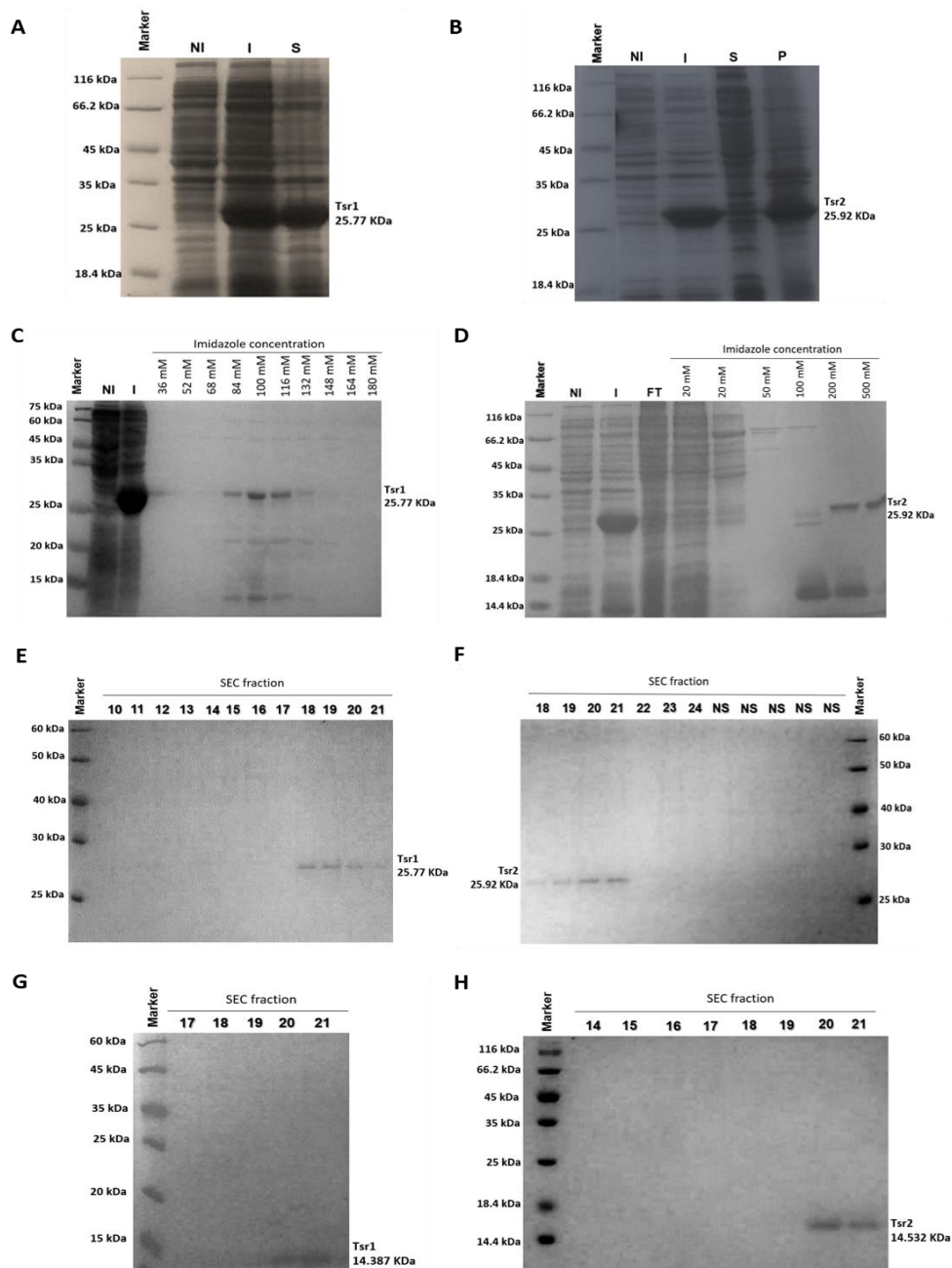
Figure 38. Expression and purification of *X. citri* Tsr1-LBD and Tsr2-LBD. **A** and **B.** Expression assays of Tsr1-LBD and Tsr2-LBD proteins. SDS-PAGE showing Tsr1-LBD (25.77 KDa) in the soluble fraction, and Tsr2-LBD (25.95 KDa) in the soluble fraction after the induction of plasmids pOP3B-Tsr1 and pOP3B-Tsr2 with 0.5 mM and 0.1 mM of IPTG, respectively. NI: Not induced, I: Induced, S: Supernatant, P: Pellet. **C** and **D**. Purification of Tsr1-LBD and Tsr2-LBD by IMAC. SDS-PAGE showing Tsr-LBD1 and Tsr2-LBD purified successfully, with some contaminants, and in low quantity. Tsr1-LBD was eluted with an imidazole concentration between 84 and 132 mM while Tsr2-LBD was eluted with an imidazole concentration between 100 and 200 mM. **E** and **F.** Purification of Tsr1-LBD and Tsr2-LBD by Size SEC. SDS-PAGE showing Tsr1-LBD and Tsr2-LBD purified without contaminants. **G** and **H.** Purification of Tsr1-LBD and Tsr2-LBD native. After GB1 cleavage from Tsr1-LBD and Tsr2-LBD proteins, proteins have the expected sizes of 14.387 KDa and 14.532 KDa, respectively.

4.2.3 The use of pUCP30T plasmid did not produce a good screening of X. citri knockout strains

To construct a *X. citri* strain with *pho*X deletion, a fragment containing *pho*X gene flanked by 500 pb upstream and downstream was amplified, with a predicted size of 2000 pb (fragment named as up-phoX-dw), and cloned into pUCP30T plasmid (resulting pUCP30T-up-phoX-dw plasmid) **(Figure 39A)**. The kanamycin gene (*km*) flanked by a FRT sites was amplified from pKD13 plasmid with primers containing 50 pb upstream and downstream *pho*X gene getting a fragment of 1404 pb (fragment named as up-km-dw) **(Figure 39B)**.

The construct was obtained after the recombination between pUCP30T-up-phoX-dw plasmid and the up-km-dw linear DNA by the Exo, Beta and Gam proteins produced from pKD46 plasmid. The insertion of the kanamycin gene was confirmed by amplification from pUCP30T-up-km-dw plasmid **(Figure 39C)** and the fragment up-km-dw was sequenced to avoid point mutations either in the upstream or downstream regions of *pho*X gene, and in the kanamycin region, flanked by FRT **(Appendix 5).**
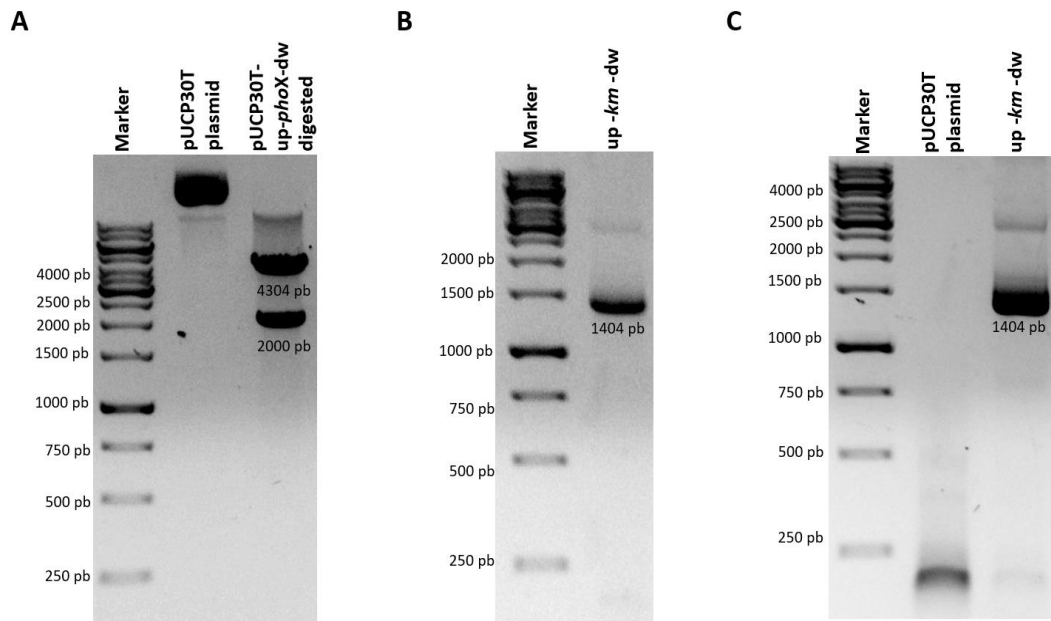
Figure 39. Agarose gel of each step performed to obtain the knockout construct (pUCP30T-up-km-dw) to delete *pho*X. A. Restriction assay of pUCP30T-up-phoX-dw plasmid with XbaI and HindIII enzymes that separated the gene up-phoX-dw of 2000 pb from pUCP30T plasmid of 4304 pb. **B.** Amplification of up-km-dw gene from pKD13 plasmid was done using primers of 70 pb containing 50 pb of the flanked genes of *pho*X, and 20 pb to amplify the *km* gene. **C.** Amplification of up-km-dw gene (using the primers in B) from pUCP30T and pUCP30T-up-km-dw plasmids, respectively.

Unfortunately, after the electroporation of *X. citri* strain with pUCP30T-up-*km*-dw plasmid, we only obtained 5 and 1 colonies in dilutions 10^0 and 10^{-1}, respectively **(Figure 40)**. Attempts to amplify of up-*km*-dw fragment failed **(Figure 41)**.
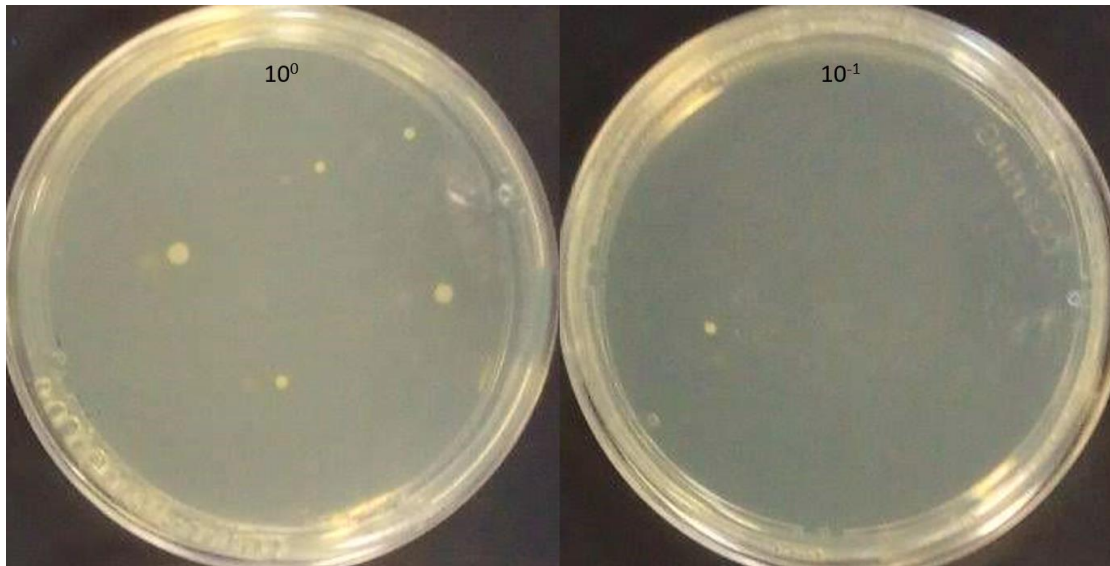
Figure 40. *X. citri* **colonies after electroporation with pUCP30T-up-*km*-dw plasmids.** Plates of dilutions 10^0 and 10^{-1} showing 5 colonies and 1 colony, respectively.
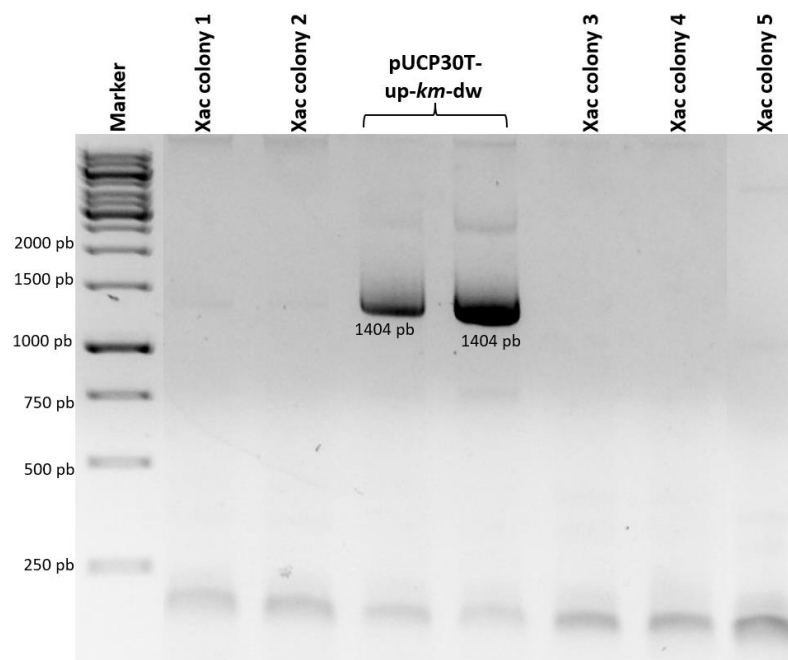
Figure 41. Colony PCR of the possible strains containing *pho*X deletion. None of the *X. citri* colonies contained the up-*km*-dw fragment.

The outcome is that pUCP30T plasmid function as a suicide vector and can generate knockout strains in *S. meliloti*, (SIBLEY *et al.,* 2006), but seems to have low efficiency in *X. citri,* since we only obtained a total of 6 colonies. Considering that in *E. coli* the recombinant frequency using the Lambda Red system is about 1 recombinant per 10^3 to 10^4 of viable cells (SWINGLE *et al.,* 2010), we needed at least a hundred colonies in the dilution 10^{-1} to have a good screening of colonies.

5 CONCLUSIONS

Altogether, the results presented in this work revealed important structural and functional features of PstS and PhoX proteins (which are PBP proteins that belong to the *X. citri* Pho regulon)*.* Firstly, we showed that besides the high amino acid sequence identity (70%) and the conservation of residues in the ligand-binding pocket, both present distinct secondary structures, specifically those that form the hinge region between domains and the loop-helix-loop in the C- Domain, which are responsible for significant differences in the protein's behavior.

The fluorescence analyses showed the first characterization of PhoX Kd to phosphate of 6.1 µM. Indeed, although the Kd for PstS has not being calculated, the fluorescence analysis and the calculations of DeltaF by bioinformatics, showed values of approx.-8.0 kcal/mol and ~7.1 kcal/mol, for PstS and PhoX, respectively, strongly suggesting that PstS has a higher affinity to phosphate than PhoX. In fact, Pegos and collaborators (2017) showed that PhoX complement the transport function in the *X. citri* carrying a *pstS* deletion, but many other differences were observed, which suggests that PstS has another function in the cell.

These differences also were revealed in the MD simulations reinforcing that despite their ability to bind phosphate, they may play unknown and different functions.

MDs of PhoX and PstS suggest a mechanism for phosphate uptake and release that would start with the movement of residues from the binding site in N- Domain that allow the first contact with the phosphate anion, and then coordination with C- Domain residues. This indicates that the flexibility of C- Domain would be important to fit the ligand inside the pocket. On the other hand, the MDs also showed the phosphate is no longer in the pocket, but goes to its entrance after destabilization of the binding site residues, with special attention to aspartic acid and arginine.

The phosphate behavior inside the binding site reveals that besides the seven residues (2 Ser, 1 Asp, 1 Arg, 1 Gly, 1 Thr and 1 Asn) identified in the crystal structure of PhoX and conserved in PstS, there are additional ones playing important function in the simulation: in PhoX, S163 and Y222, and in PstS, G186 and S187.

We proposed two hypotheses for PBPs of Pst system in *X. citri*: (i) The presence of lysine, K70 and K199 in PhoX, and K94 and K223 in PstS, which are located in the region of interaction with permeases PstA and PstC, could be key residues for phosphate delivery into permeases; (ii) These lysines could have a function of sensing, generating a signal of phosphate starvation. These two hypotheses can be further explored to clarify and answer how PBPs sense and help in the phosphate transportation inside the permeases.

Unfortunately, although the putative Tsr proteins from *X. citri* were expressed and purified, it was not possible to proceed with their characterization. Further studies need to be performed to confirm their participation in *X. citri* chemotaxis.

REFERENCES

AN, S. et al. Mechanistic insights into host adaptation, virulence and epidemiology of the phytopathogen *Xanthomonas*. **FEMS Microbiology Reviews**, v. 44, n. 1, 2020.

BALAN, A. et al. Purification and in vitro characterization of the maltose-binding protein of the plant pathogen *Xanthomonas citri*. **Protein Expr Purif,** v.43, n.2, p.103-110, 2005.

BARTUAL, S. et al. Three-dimensional structures of Lipoproteins from *Streptococcus pneumoniae* and *Staphylococcus aureus*. **International Journal of Medical Microbiology**, v. 308, p. 692-704, 2018.

BERNTSSON, R. et al. A structural classification of substrate-binding proteins. **FEBS Letters,** v. 584, n. 12 p. 2606-2617, 2010.

BI, S.; LAI, L. Bacterial chemoreceptors and chemoeffectors. **Cell. Mol. Life Sci**, v. 72**,** p. 691–708, 2015.

BLANCO, A. G. et al. The structure of a transcription activation subcomplex reveals how σ (70) is recruited to PhoB promoters. **EMBO J,** v. 30, n. 18, p.76-85, 2011.

BOCK, C.; GOTTWALD, T.; PARKER, P. Distribution of canker lesions on the surface of diseased grapefruit. **Plant Pathology,** v. 60, p. 986–991, 2011.

BEEK, J.; GUSKOV, A.; SLOTBOOM, D. Structural diversity of ABC transporters. **The Rockefeller University Press J. Gen. Physiol,** v.143, n. 4, p. 419–435, 2014.

BERG J.; TYMOCZKO J.; STRYER L. Biochemistry. 5th edition. New York: W H Freeman; Chapter 3 **Protein Structure and Function,** 2002.

CAICEDO, J.; VILLAMIZAR, S. *Xanthomonas citri* ssp. *citri* Pathogenicity, a Review. **Research, Development and Biotechnology,** p. 1-11, 2021.

CATARA, V. et al. Trends in Molecular Diagnosis and Diversity Studies for Phytosanitary Regulated *Xanthomonas*. **Microorganisms,** v. 9, p. 862, 2021.

CARTAILLER, J.; LUECKE, H. Structural and functional characterization of pi bulges and other short intrahelical deformations. **Structure,** v. 12, p. 133-44, 2004.

COX, G.; WEBB, D.; ROSENBERG, H. Specific Amino Acid Residues in Both the PstB and PstC Proteins Are Required for Phosphate Transport by the *Escherichia coli* Pst System. **Journal of Bacteriology,** v.3, n. 1, p. 1531-1534, 1989.

CREPIN, S. et al. The Pho regulon and the pathogenesis of *Escherichia coli.* **Veterinary Microbiology,** v. 153, p. 82–88. 2011.

DE SAINT-EXUPERY, A. **The Little Prince.** Translated by Irene Testot-Ferry, English Editions, 2018.

DAS, A.K. Citrus canker- A review. **J. Appl. Hort,** v. 5, n. 1, p. 52-60, 2003.

DE BOER, M. et al. Conformational and dynamic plasticity in substrate-binding proteins underlies selective transport in ABC importers. **Elife.** v. 22, n. 8, 2019.

DHAKAL, D.; REGMI, C.; BASNYAT, S. R. Etiology and Control of Citrus Canker Disease in Kavre. **Nepal Journal of Science and Technology,** v. 10, p. 57-61, 2009.

DÍAZ, M. et al. THE high-affinity phosphate-binding protein PstS is accumulated under high fructose concentrations and mutation of the corresponding gene affects differentiation in *Streptomyces lividans.* **Microbiology,** v. 151, p. 2583–2592, 2005.

DWYER, M.; HELLINGA, H. Periplasmic binding proteins: a versatile superfamily for protein engineering. **Curr Opin Struct Biol,** v. 14, n. 4, p. 495-504, 2004.

EDWARDS, K. Periplasmic-binding protein-based biosensors and bioanalytical assay platforms: Advances, considerations, and strategies for optimal utility. **Talanta Open,** v. 3, 2021.

ELIAS, M.; WELLNER, A.; GOLDIN-AZULAY, K*.* The molecular basis of phosphate discrimination in arsenate-rich environments. **Nature,** v. 491, p. 134–137, 2012.

ERIKSSON, S.; NORDEN, B.; TAKAHASHI, M. Binding of DNA Quenches Tyrosine Fluorescence of RecA without Energy Transfer to DNA Bases. **The Journal of Biological Chemistry,** v. 268. n. 3, p. 1805-1810, 1993.

FONSECA, N. et al. Detection and identification of *Xanthomonas* pathotypes associated with citrus diseases using comparative genomics and multiplex PCR**. PeerJ,** p.01-22, 2019.

FUKAMI-KOBAYASHI, K.; TATENO, Y.; NISHIKAWA, K. Domain dislocation: a Change of Core Structure in Periplasmic Binding Proteins in their Evolutionary History. **J. Mol. Biol,** v. 286, p. 279-290, 1999.

FURTWÄNGLER, K. et al. Regulation of phosphate uptake via Pst transporters in *Halobacterium salinarum* R1. **Molecular Microbiology,** v.76, n. 2, p. 378–392, 2010.

GARDNER, S.; MCCLEARY, W. Control of the phoBR Regulon in *Escherichia coli.* **EcoSal Plus.** 2019.

GHISAIDOOBE, A.; CHUNG, S. Intrinsic tryptophan fluorescence in the detection and analysis of proteins: a focus on Förster resonance energy transfer techniques. **International journal of molecular sciences**, v. 15, n. 12, p. 22518–22538, 2014.

GOCHEZ, A. et al. Pacbio sequencing of copper-tolerant *Xanthomonas citri* reveals presence of a chimeric plasmid structure and provides insights into reassortment and shuffling of transcription activator-like effectors among *X. citri* strains. **BMC Genomics**, v.19, p. 16, 2018.

GONZALEZ, D.; RICHEZ, M.; BERGONZI, C. Crystal structure of the phosphate-binding protein (PBP-1) of an ABC-type phosphate transporter from *Clostridium perfringens*. **Sci Rep,** v. 4, p. 636, 2014.

GOTTIG, N. et al. Mechanisms of infection used by *Xanthomonas axonopodis* pv. c*itri* in citrus canker disease. **Current Research, Technology and Education Topics in Applied Microbiology and Microbial Biotechnology,** p. 196-204, 2010.

GRAHAM, J. et al. *Xanthomonas axonopodis* pv. *citri*: factors affecting successful eradication of citrus canker. **Molecular Plant Pathology**, v. 5, n. 16, p.1-15, 2004.

GREENFIELD, N. Using circular dichroism spectra to estimate protein secondary structure. **Nat Protoc,** v. 1, p. 2876–2890, 2006.

GOTTWALD, T. R.; GRAHAM, J. H.; SCHUBERT, T. S. Citrus canker: The pathogen and its impact. Online. **Plant Health Progress**. 2002.

HADDAD A. Environmental factors influence virulence of *Pseudomonas aeruginosa.* Von der Fakultät für Lebenswissenschaften der Technischen **Universität Carolo-Wilhelmina,** 2010.

HOLLENSTEIN, K.; FREI, D.; LOCHER, K. Structure of an ABC transporter in complex with its binding protein. **Nature**, v. 446, n. 7132, p. 213-216, 2007.

JACOBSEN, S. et al. The high-affinity phosphate transporter Pst is avirulence factor for *Proteus mirabilis* during complicated urinary tract infection. **FEMS Immunol Med Microbiol,** n. 52, p. 180–193, 2008.

KIEL, C.; YUS, E.; SERRANO, L. Engineering Signal Transduction Pathways. **Cell 140**. **Elsevier Inc,** 2010.

KURODA, A. et al. Evaluation of Phosphate Removal from Water by Immobilized Phosphate-Binding Protein PstS. **Journal of Bioscienci and Bioengineering,** v. 90, n. 6, p. 688-690, 2000.

LAMARCHE, M. et al. The phosphate regulon and bacterial virulence: a regulatory network connecting phosphate homeostasis and pathogenesis. **FEMS Microbiol Rev,** v. 32, p. 461-473, 2008.

LEDVINA, P. et al. Dominant role of local dipolar interactions in phosphate binding to a receptor cleft with an electronegative charge surface: equilibrium, kinetic, and crystallographic studies. **Protein Sci,** v. 12, p. 2550, 1998.

LIU, C. et al. Two-Component Signal Transduction Systems: A Major Strategy for Connecting Input Stimuli to Biofilm Formation. **Front. Microbiol,** n. 9, 2019.

LI, H. et al. Ligand-induced structural changes analysis of ribose-binding protein as studied by molecular dynamics simulations. **Technol Health Care,** v. 29, n. S1, p. 103-114, 2021.

LOCHER, K. P. Mechanistic diversity in ATP-binding cassette (ABC) transporters. N**ature Structural & Molecular Biology,** v. 23, n. 6, 2016.

MARCIO, L. et al. Chemotactic signal transduction and phosphate metabolism as adaptive strategies during citrus canker inductionby *Xanthomonas citri*, **Springer. Funct Integr Genomics,** v.18, 2014.

MARTÍN, J.; LIRAS, P. Molecular Mechanisms of Phosphate Sensing, Transport and Signalling in *Streptomyces* and Related Actinobacteria. **Int. J. Mol. Sci**, 2021.

MERTENS, H.; SVERGUN, D. Structural characterization of proteins and complexes using small-angle X-ray solution scattering. **J Struct Biol,** v. 172, n. 1, p. 128-41, 2010.

MCCLEARY, W. Molecular Mechanisms of Phosphate Homeostasis in *Escherichia coli*. **Intench,** Chapter 17. p. 334- 357, 2017.

MICHIGAMI, T. et al. Phosphate as a signaling molecule and its sensing mechanism. **Physiol Rev,** n. 98, p. 2317-2348, 2018.

MICSONAI, A. et al. Accurate secondary structure prediction for CD. **Proceedings of the National Academy of Sciences,** v. 112, n. 24, p. E3095-E31303, 2015.

MITROPHANOV, A. et al. Signal integration in bacterial two-component regulatory systems. **Genes & Development by Cold Spring Harbor Laboratory Press,** n. 22, p. 2601–2611, 2008.

MONEDERO, V.; REVILLA-GUARINOS, A.; ZUNIGA, M. Physiological Role of Two-Component Signal Transduction Systems in Food-Associated Lactic Acid Bacteria. **Advances in Applied Microbiology,** v. 99, 2017.

MOWERY, P. et al. Chemotactic Signaling by Single-Chain Chemoreceptors. **PLoS ONE**, v.10, n. 12, 2015.

NAYEM, S.A. et al. Phytohormone- Induced Resistance against *Xanthomonas axonopodis* PV. *citri* in *Citrus aurantifolia*. **American Journal of Plant Sciences**, v. 8, p. 1135-1147, 2017.

PATANÉ, J.S. et al. Origin and diversification of *Xanthomonas citri* subsp. *citri* pathotypes revealed by inclusive phylogenomic, dating, and biogeographic analyses. **BMC Genomics**, v.20, n. 700, 2019.

PEGOS, V.; MEDRANO, F.; BALAN, A. Crystallization and preliminary X-ray diffraction analysis of the phosphate-binding protein PhoX from *Xanthomonas citri*. **Acta Cryst,** n. 70, p. 1604–1607, 2014.

PEGOS, V. et al. Phosphate regulated proteins of *Xanthomonas citri* subsp. *citri*: A proteomic approach. **Journal of Proteomics**, v. 108, p. 78–88, 2014.

PEGOS, V. et al. Structural features of PhoX, one of the phosphate-binding proteins from Pho regulon of *Xanthomonas citri*. **PLOS ONE,** v. 12, n. 5, 2017.

PEGOS V. Structural an Funtional studies of the Phosphate sysema uptake from bactéria phytopathogenic *Xanthomonas axonopodis* pv. *citri*. **University of Campinas,** 2015.

PARKINSON, J.; HAZELBAUER, G.; FALKE, J. Signaling and sensory adaptation in *Escherichia coli* chemoreceptors: 2015 update. **Trends Microbiol,** v. 23, n. 5, p. 257-66, 2015.

RAMBO, R.; TAINER, J. Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law. **Biopolymers,** v. 95, n. 8, p. 59-71, 2011.

RAVI, B.; YOGAVEL, C.; SHARMA, A. Structural Analysis of ABC-family Periplasmic Zinc Binding Protein Provides New Insights Into Mechanism of Ligand Uptake and Release. **J Mol Biol,** v. 6, n. 367, p. 970–982, 2007.

REINARTZ, I. ET AL. Structural Analysis of a Genetically Encoded FRET Biosensor by SAXS and MD Simulations. **Sensors,** v.21, n.4144, 2021.

RICHARD, D. et al. Genome Sequences of Six Copper-Resistant *Xanthomonas citri* pv. *citri* Strains Causing Asiatic Citrus Canker, Obtained Using Long-Read Technology, **Genome announcements,** v. 5, n. 12, 2017.

RICO-JIMÉNEZ, M. et al. Two different mechanisms mediate chemotaxis to inorganic phosphate in *Pseudomonas aeruginosa.* **Scientific Reports,** v.6, 2016.

RIGHETTI, E.; ULUSEKER, C.; KAHRAMANOGULLARI, O. Simulations as a Tool for Assessing Signal Fidelity in Gene Expression in Synthetic Promoter Design**. Biology,** v.10, p.724. 2021.

RIZK, S.; CUNEO, M.; HELLINGA, H. Identification of cognate ligands for the *Escherichia coli* phnD protein product and engineering of a reagentless fluorescent biosensor for phosphonates. **Protein Sci,** v.15, n. 7, p.1745-1751, 2006.

RYAN, R. et al. Pathogenomics of *Xanthomonas*: understanding bacterium–plant interactions. **Nature reviews. Microbiology,** v. 9, p. 344-355, 2011.

SALAH, U.; ROUJEINIKOVA, A. Methyl-accepting chemotaxis proteins: a core sensing element in prokaryotes and archaea. **Cell. Mol. Life Sci***, v. 74, p. 3293–3303, 2017.

SANTOS-BENEIT, F.; RODRIGUEZ-GARCIA, A.; FRANCO-DOMINGUEZ, E. Phosphate-dependent regulation of the low- and high-affinity transport systems in the model actinomycete *Streptomyces coelicolor.* **Microbiology,** n. 154, p. 2356–2370, 2008.

SANTOS-BENEIT, F. The Pho regulon: a huge regulatory network in bacteria. **Front. Microbiol,** v. 6, n. 402, 2015.

SCHRAMKE, H.; et al. Revisiting regulation of potassium homeostasis in *Escherichia coli*: the connection to phosphate limitation. **Microbiology Open,** p. 1–16, 2017

SCHEEPERS, G. et al. An updated structural classification of substrate-binding proteins. **FEBS Lett,** v.590, n. 23, p. 4393-4401, 2016.

SCHUBERT, T. et al. Meeting the Challenge of Eradicating Citrus Canker in Florida- Again. **Plant Disease**, v.85, n. 4, 2001.

SHARMA, A.; GAUTAM, S.; WADHAWAN, S. *Xanthomonas.* **Bhabha Atomic Research Centre. Elsevier,** 2014.

SHARMA, S.; SHARMA, R. Citrus Canker Approaching Century: A Review. **Tree and Forestry Science and Biotechnology**, v. 3, n. 2, p. 54-65, 2009.

SHEVCHUK, R.; HUB, J. Bayesian refinement of protein structures and ensembles against SAXS data using molecular dynamics. **PLoS Comput Biol,** v.13, n.10, 2017.

SHRESTHA, S. et al. Rationally designed fluorescently labeled sulfate-binding protein mutants: evaluation in the development of a sensing system for sulfate. **Biotechnol Bioeng.** v. 78, n.5, p. 517-26, 2002.

SIBLEY, C.; MACLELLAN, S.; FINAN, T. The *Sinorhizobium meliloti* chromosomal origin of replication. **Microbiology**, v.152, p.443–455, 2006.

SINDREWICZ, P.; LI, X., YATES, E.A. Intrinsic tryptophan fluorescence spectroscopy reliably determines galectin-ligand interactions. **Sci Rep,** v. 9, 2019.

SOARES, M. et al. Proteome of the phytopathogen *Xanthomonas citri* subsp. *citri*: a global expression profile. **Proteome Science,** v.8, n. 55, 2010.

SOLSCHEID, C. et al. Development of a Reagentless Biosensor for Inorganic Phosphate, Applicable over a Wide Concentration Range. **Biochemistry,** v. 54, p. 5054 – 5062, 2015.

SPURLINO, J.; LU, G.; QUIOCHO, F. The 2.3 A resolution structure of the maltose- or maltodextrin-binding protein, a primary receptor of bacterial active transport and chemotaxis. **J Biol Chem**, v.266, n.8, p.5202-5219. 1991.

STOCK, J.; BAKER, M. Physiology. Chemotaxis. **Elsevier Inc**, 2009.

STOCKNER, T.; VOGEL, H.; TIELEMAN, D. A salt-bridge motif involved in ligand binding and large-scale domain motions of the maltose-binding protein. **Biophys J,** v.89, n.5, p.3362-3371, 2005.

SWINGLE, B. et al. Oligonucleotide recombination in Gram-negative bacteria. **Mol Microbiol**, v. 75, n. 1, p. 138-148, 2010.

TANG, J. et al. *Xanthomonas campestris* Pathovars. **Trends in Microbiology. Cells Press Reviews,** 2020.

TAVARES, D. et al. Computational redesign of the *Escherichia coli* ribose-binding protein ligand binding pocket for 1,3-cyclohexanediol and cyclohexanol. **Scientifics Reports,** v.9, n.16940, 2019.

TETSCH L.; JUNG K. The regulatory interplay between membrane-integrated sensors and transport proteins in bacteria. **Molecular Microbiology,** v. 73, n. 6, p. 982–991, 2009.

TIMILSINA, S. et al. *Xanthomonas* diversity, virulence and plant–pathogen interactions. **Nature Reviews, Microbiology.** v. 18, 2020.

TIWARI, S. et al. Two-Component Signal Transduction Systems of Pathogenic Bacteria as Targets for Antimicrobial Therapy: An Overview. **Front. Microbiol,** n. 8, 2017.

THOMAS, C. et al. Structural and functional diversity calls for a new classification of ABC transporters. **FEBS Lett,** n. 594, p. 3767-3775, 2020.

THOMAS, C.; TAMPÉ, R. Annual Review of Biochemistry Structural and Mechanistic Principles of ABC Transporters. **Annu. Rev. Biochem,** v. 89, 2020.

TORRES, M.; FORMAN H.; J. Signal transduction. **Elsevier Ltd**, 2006.

ULUŞEKER, C. et al. Quantifying dynamic mechanisms of auto-regulation in *Escherichia coli* with synthetic promoter in response to varying external phosphate levels. **Scientific Reports,** v. 9, p. 2076, 2019.

VUPPADA, R. et al. Phosphate signaling through alternate conformations of the PstSCAB phosphate transporter. **BMC Microbiology,** v.18, n.8, 2018.

WILKENS, S. Structure and mechanism of ABC transporters. **F1000Prime Reports,** v. 7, p. 14, 2015.

YAMMINE, A.; GAO, J.; KWAN, A. Tryptophan Fluorescence Quenching Assays for Measuring Protein-ligand Binding Affinities: Principles and a Practical Guide. **Bio-protocol,** v.9, n.11, 2019.
YAO, N. et al. Modulation of a salt link does not affect binding of phosphate to its specific active transport receptor. **Biochemistry,** v. 7, n. 14, p. 2079-85, 1996.

APPENDICES

Appendix 1. Maps of pET28a⁺-PhoX and pET28a⁺-PstS plasmids designed in SnapGene program**. A.** Design of pET28a⁺-PhoX plasmid showing that this plasmid expresses *X. citri* PhoX protein without the 24 first residues and with a 6 His Tag in the N-terminal**. B**. Design of pET28a+-PstS plasmid showing that this plasmid expresses *X. citri* PstS protein without the 20 first residues and with a 6 His Tag in the N-terminal.
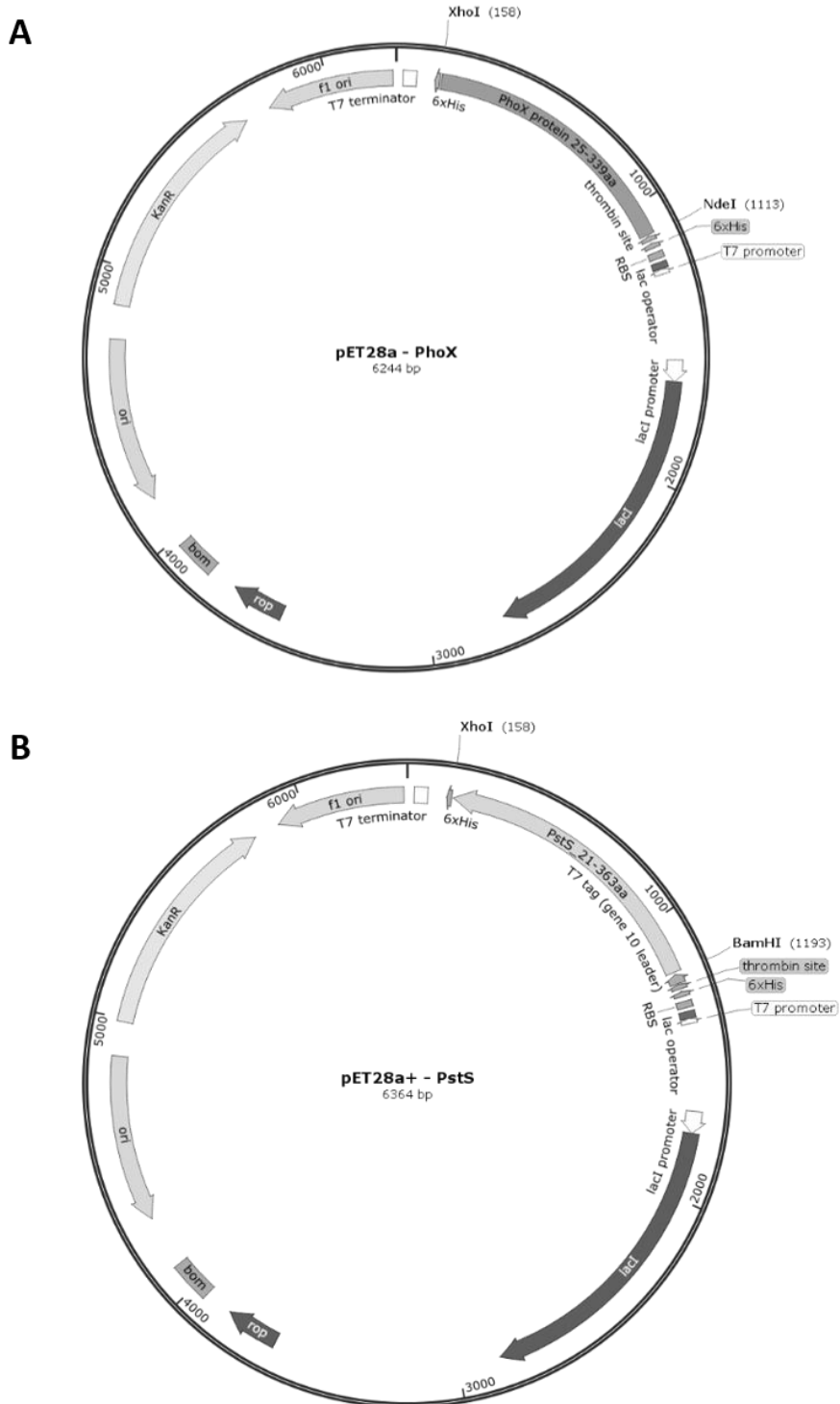
Appendix 2. Maps of pOP3BP-Tsr1 and pOP3BP-Tsr2 plasmids designed in SnapGene program. **A.** Design of pOP3BP-Tsr1 plasmid showing that this plasmid expresses the *X. citri* LBD of Tsr1 with the fusion protein GB1 and an 8 His Tag in the N-terminal. **B.** Design of pOP3BP-Tsr2 plasmid showing that this plasmid express *X. citri* LBD of Tsr2 with the fusion protein GB1 and an 8 His Tag in the N-terminal.

A. pOP3BP - Tsr1, 5046 bp

B. pOP3BP - Tsr2, 5052 bp
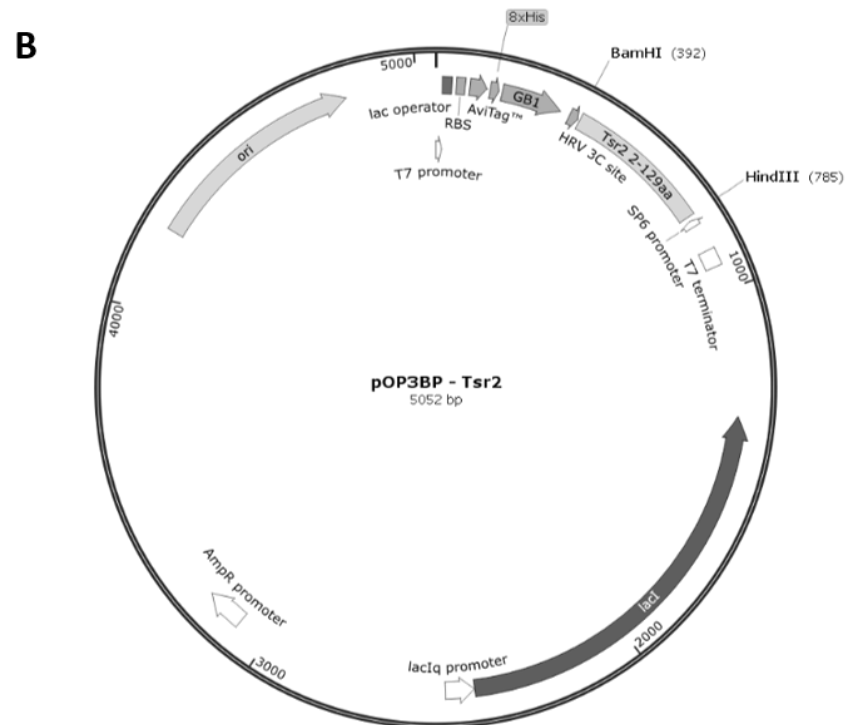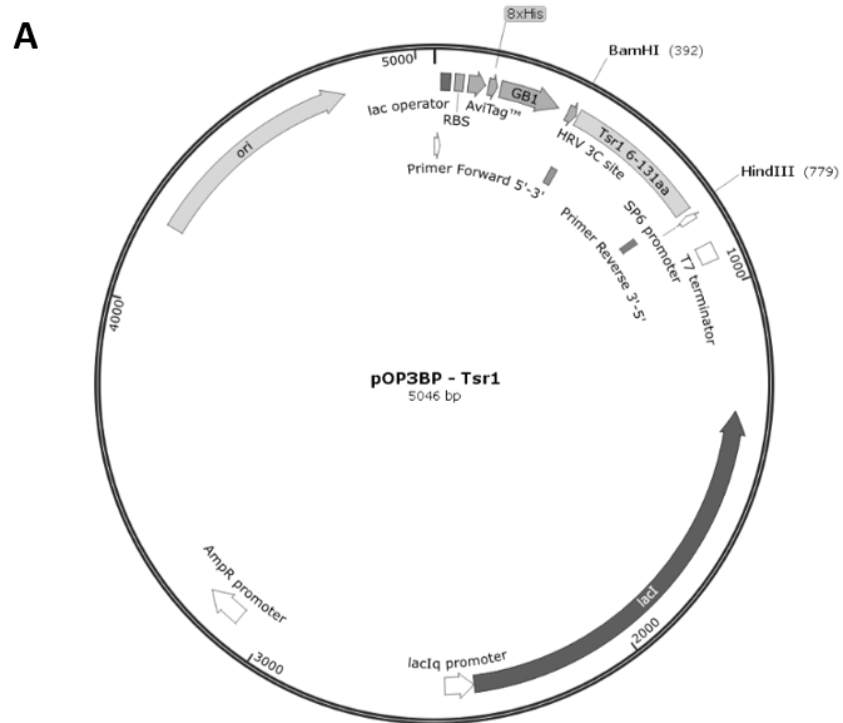
Appendix 3. Maps of pKD46 and pKD13 plasmids designed in SnapGene program. **A.** Design of pKD46 plasmid containing the genes that encodes proteins Gam, Beta and Exo responsible for the process of homologous recombination. **B.** Design of pKD13 plasmid containing the gene that confers resistance to neomycin and kanamycin antibiotics, gene used to produce the knockout plasmid for *X. citri phoX* gene (pUCP30T-up-km-dw).
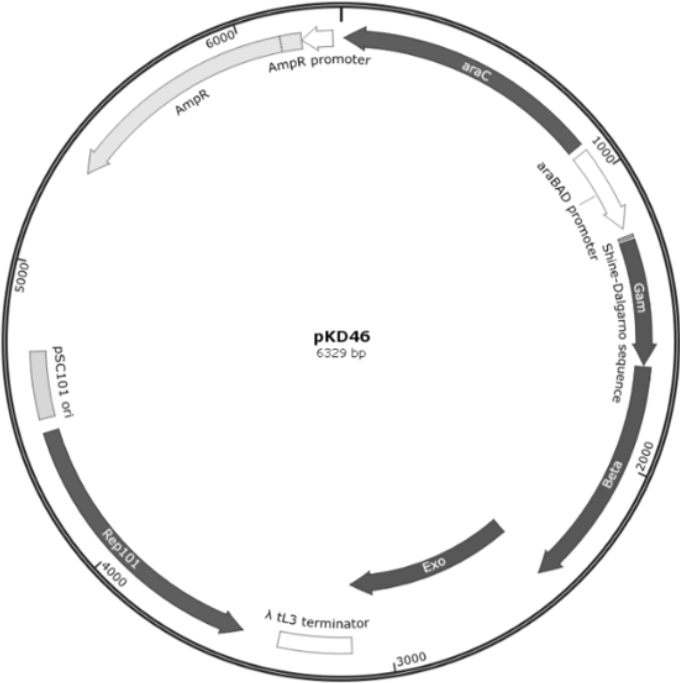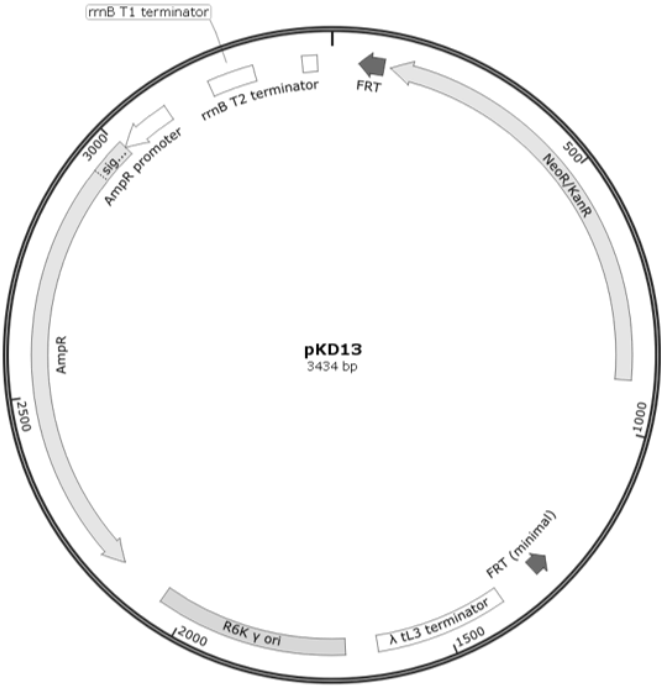
Appendix 4. Ramachandran plot of the generated PstS model. The graphic shows the phi-psi torsion angles for all residues of the PstS model. The black squares represent the allowed phi-psi backbone, and the red areas corresponds to the most favorable combinations of phi-psi values.

PROCHECK

Ramachandran Plot
3568706

Psi (degrees)

Phi (degrees)
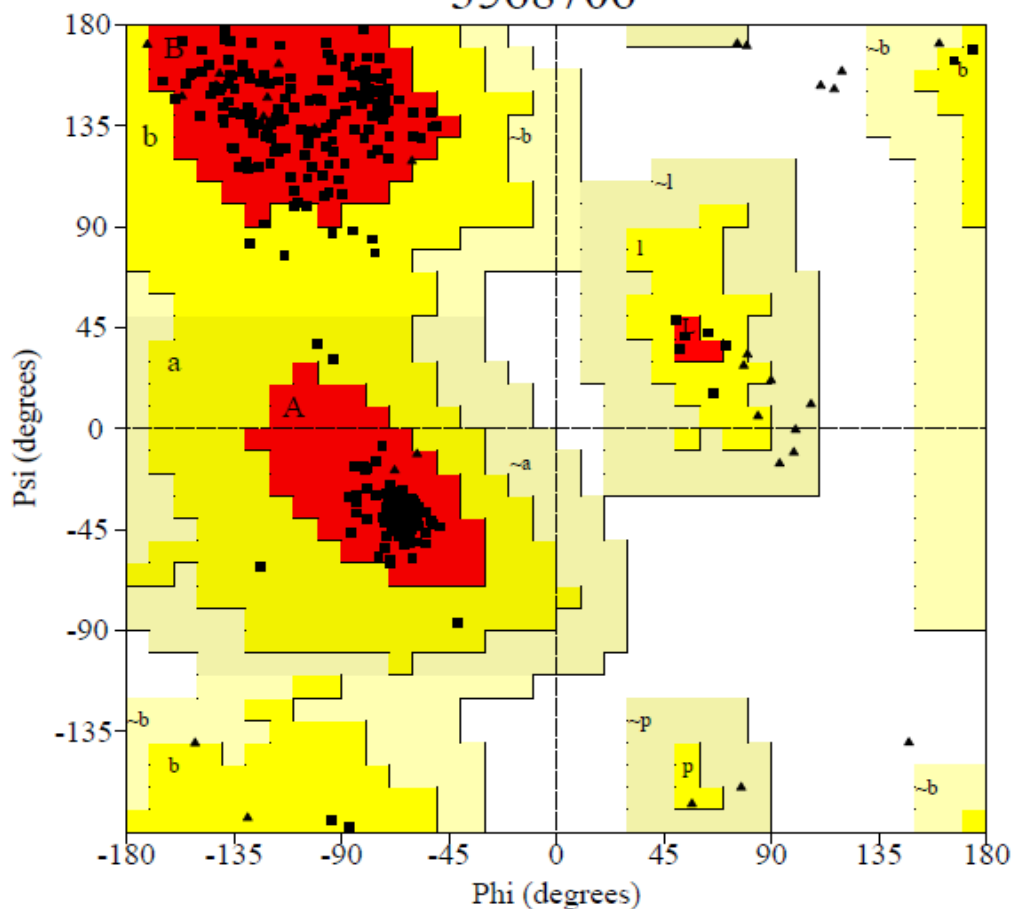
Plot statistics

Residues in most favoured regions [A,B,L]	285	94.1%
Residues in additional allowed regions [a,b,l,p]	18	5.9%
Residues in generously allowed regions [~a,~b,~l,~p]	0	0.0%
Residues in disallowed regions	0	0.0%
	----	------
Number of non-glycine and non-proline residues	303	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	39	
Number of proline residues	19	

Total number of residues	363	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

Appendix 5. Sequencing of *up-km-dw* gene from pUCP30T-*up-km-dw* knockout plasmid. The *up-km-dw* gene consist of three parts: 500 pb upstream *phoX* gene, kanamycin resistant gene and 500 pb downstream *phoX* gene. **A.** Map of *up-km-dw* gene showing its three parts**. B.** Nucleotide sequences of the 500 pb upstream *phoX* gene plus part of the kanamycin gen**. C.** Nucleotide sequences of the 500 pb downstream *phoX* gene plus part of the kanamycin gen.

A

up-km-dw gene
2337 bp

B

Upstream *pho*X gene

GCAGCACGCGCGGCCTGGTCGACGACAATCCGTCGATCATCGAAGCACGCGCGCAGTTTTACTGGTAAGTGCGTCGCTGCCGCGTGCGCTCGG

273   283   293   303   313   323   333   343   353   363

Upstream *pho*X gene

GGTCGAGGCACGCGGTGGTAGCCGCCACTTCCTGCCGGCCCGAACATCCGACGGCGCAACGCCGTCATCGGCCTGTCATGTCCCAGACACCG

363   373   383   393   403   413   423   433   443   453

Upstream *pho*X gene

CGCATGGTCATGCCGTTGCCCTGCAATGGCACCACGGCGGGACGTCCTCATCCCTTCGCTACAGGAGCTGTTCCGTGTAGGCTGGAGCTGC

453   463   473   483   493   503   513   523   533   543

Kanamycin resistance gene

GCTTCGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTTCAAGATCCCCTTATTAGAAGAACTCGTCAAGAAGGCGATA

543   553   563   573   583   593   603   613   623   633

Kanamycin resistance gene

AGAAGGCGATGCGCTGCGAATCGGGAGCGGCGATACCGTAAAGCACGAGGAAGCGGTCAGCCCATTCGCCGCCAAGCTCTTCAGCAATATCA
633　643　653　663　673　683　693　703　713　723

Kanamycin resistance gene

CACGGGTAGCCAACGCTATGTCCTGATAGCGGTCCGCCACACCCAGCCGGCCACAGTCGATGAATCCAGAAAAGCGGCCATTTTCCACCATG
723　733　743　753　763　773　783　793　803　813

Kanamycin resistance gene

CGCATGGTCATGCCGTTGCCCTGCAATGGCACCACGGCGGGACGTCCTCATCCCTTCGCTACAGGAGCTGTTCCGTGTAGGCTGGAGCTGC
813　823　833　843　853　863　873　883　893　903

Kanamycin resistance gene

CTTCGAAGTTCCTATACTTTCTAGAGAATAGGAACTTCGGAATAGGAACTTCAAGATCCCCTTATTAGAAGAACTCGTCAAGAAGGCGATA
903　913　923　933　943　953　963　973　983　993

Kanamycin resistance gene

A G A A G G C G A T G C G C T G C G A A T C G G G A G C G G C G A T A C C G T A A A G C A C G A G G A A G C G G T C A G C C C A T T C G C C G C C A A G C T C T T C A G C A A T A T C A

993 1003 1013 1023 1033 1043 1053 1063 1073 1083

C

Kanamycin resistance gene

C T G A C A G C C G G A A C A C G G C G G C A T C A G A G C A G C C G A T T G T C T G T T G T G C C A G T C A T A G C C G A A T A G C C T C T C C A C C C A A G C G G C C G G A G A

1281 1291 1301 1311 1321 1331 1341 1351 1361 1371

Kanamycin resistance gene

A A C C T G C G T G C A A T C C A T C T T G T T C A A T C A T G C G A A A C G A T C C T C A T C C T G T C T C T T G A T C A G A T C T T G A T C C C C T G C G C C A T C A G A T C C T

1371 1381 1391 1401 1411 1421 1431 1441 1451 1461

T T G G C G G C A A G A A A G C C A T C C A G T T T A C T T T G C A G G G C T T C C C A A C C T T A C C A G A G G G C G C C C C A G C T G G C A A T T C C G G T T C G C T T G C T G T

1461 1471 1481 1491 1501 1511 1521 1531 1541 1551

T C C A T A A A A C C G C C C A G T C T A G C T A T C G C C A T G T A A G C C C A C T G C A A G C T A C C T G C T T T C T C T T T G C G C T T G C G T T T T C C C T T G T C C A G A T

1551   1561   1571   1581   1591   1601   1611   1621   1631   1641

A G C C C A G T A G C T G A C A T T C A T C C G G G G T C A G C A C C G T T T C T G C G G A C T G G C T T T C T A C G T G T T C C G C T T C C T T T A G C A G C C C T T G C G C C C T

1642   1651   1661   1671   1681   1691   1701   1711   1721   1731

T G A G T G C T T G C G G C A G C G T G A G C T T C A A A A G C G C T C T G A A G T T C C T A T A C T T T C T A G A G A A T A G G A A C T T C G A A C T G C A G G T C G A C G G A T

1732   1741   1751   1761   1771   1781   1791   1801   1811   1821

Downstream *phoX* gene

T C C C C G G A A T G A G T A C G G C G C A G G C C T C T G C T G C A G C G G C C T G G T G G G G C A A G C C C A T G C T C T A A A A C G G C A T G G G C C G T T G T C T G G T C C A T G

1821   1831   1841   1851   1861   1871   1881   1891   1901   1911

Downstream *phoX* gene

TGCCGGCGAGCCGCCCGCCCACCTGGAACTGGCAATGGGCGCCGTGCACTAGGGCCTGGCAGCAAGCACGAAAACGCGTGAAGGAACGATGC
1911   1921   1931   1941   1951   1961   1971   1981   1991   2001

Downstream *phoX* gene

GCCGTCTCTCTCCCCGGCGTCGGCCCCGGGACGCGGATCCCGGAGGGTGCAACGCCCTCCGTGGTCCGCGTCTTTTTTGTGATGCGTTGGC
2001   2011   2021   2031   2041   2051   2061   2071   2081   2091

Downstream *phoX* gene

CAGAAGGCGCTGTCATGCGGGTGTAACAAAAACATCATTTCATAGTGTGGTCTGCCGGAGACACCGGCCAATCCAGCCTCGGAGCTTCCAT
2091   2101   2111   2121   2131   2141   2151   2161   2171   2181

Downstream *phoX* gene

ATGAAATTGCAGTCGGCCAGCCTGACCGCCCTGTCCCTCACCATTGCCCTCGCCCTGGCCGCGTGCTCGCCCGGCAAGGACGCGCAGTCCGG
2181   2191   2201   2211   2221   2231   2241   2251   2261   2271

Downstream *phoX* gene

GCAGACGCCGCAAAGG GCGC GCCA GCCGCAGGC GCC GCC GCCAGCG ACAGCA AGGG TG CTGAAATT T

2271   2281   2291   2301   2311   2321   2331   2337