GMR

A *Xanthomonas citri* subsp *citri* hypothetical protein related to virulence contains a non-functional HD domain and is implicated in flagellar motility

F.C.F. Vieira[1], A.M. Gonçalves[1], E.F.R. Mendoza[1,2], R.M. Ferreira[1], M.L.M. Costa[1], T.S. Balbuena[1], H.G. Sebinelli[2], P. Ciancaglini[2], J.M. Pizauro Junior[1] and J.A. Ferro[1]

[1]Departamento de Tecnologia, Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista, Jaboticabal, SP, Brasil
[2]Departamento de Química, Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, SP, Brasil

Corresponding author: J.A. Ferro
E-mail: jesus@fcav.unesp.br

ABSTRACT. Citrus canker, caused by the Gram-negative bacterium *Xanthomonas citri* subsp *citri* (Xac), severely affects most economically important citrus varieties worldwide. A previous study showed that disruption of the ORF XAC1201 from the Xac 306 strain by transposon Tn5 decreased bacterium virulence in the Rangpur lime host (*Citrus limonia* L. Osbeck). However, little is known regarding the possible function of the hypothetical protein XAC1201 and how it affects the virulence of Xac 306. Here, we confirmed that disruption of ORF XAC1201 reduces Xac 306 virulence in two different hosts, delaying the onset of typical symptoms. *In silico* analysis suggested that XAC1201

interacts with the flagellar proteins FliM and FliL, known to be an important factor for virulence. In fact, motility assays revealed that the XAC1201 mutant has a significant difference in motility compared to the wild-type Xac 306. Also, a 3-D structure model revealed modified cofactor binding sites and suggested that XAC1201 has a non-functional HD domain. This hypothesis was confirmed by enzymatic assays performed in purified, XAC1201 recombinant protein expressed in *Escherichia coli*, which revealed no significant activities previously associated with HD domains for the tested substrates. Thus, the role of the XAC1201 protein in Xac 306 virulence seems to be related to flagellar motility, although a non-classic role for the HD domain cannot be dismissed.

Key words: Citrus canker; ORF XAC1201; Motility; HD domain; Bacterial flagella

INTRODUCTION

Citrus canker, caused by the Gram-negative bacterium *Xanthomonas citri* subsp *citri* (Xac), is one of the most important citrus diseases, severely affecting most economically important citrus varieties worldwide (Das, 2003) with no efficient way of controlling the disease currently available (Brunings and Gabriel, 2003; Graham et al., 2004). The genome of *X. citri* subsp *citri* strain 306 (Xac 306) was sequenced over a decade ago (da Silva et al., 2002) and functional genomic approaches have widened our knowledge of Xac-citrus interaction and remain the most promising way to find targets for the development of treatment strategies.

A previous study showed that disruption of the ORF XAC1201 from Xac 306 by transposon Tn5 decreased bacterium virulence in the Rangpur lime host (*Citrus limonia* L. Osbeck) (Laia et al., 2009). However, little is known regarding its structure and function. XAC1201 (AAM36073.1) is a hypothetical protein containing 270 amino acids and a conserved HD superfamily domain. As HD domain architecture varies greatly, resulting in a large variety of functions, usually involving metal-dependent phosphohydrolase activity (Aravind and Koonin, 1998), its precise role, if any, in Xac 306 virulence must be investigated. Thus, to shed light on the possible function of the XAC1201 protein, we used the recombinant XAC1201 protein produced in *Escherichia coli* to assay the enzymatic functionality of its HD domain and also performed *in silico* analysis of the protein sequence.

MATERIAL AND METHODS

Bacterial growth conditions

The wild-type *X. citri* subsp *citri* strain 306 (Xac 306) was provided by IAPAR-UEL, Londrina, Brazil, and the XAC1201 mutant was obtained from a mutant library, previously described by Laia et al. (2009). EZ-Tn5 transposon insertion of the XAC1201 mutant occurred between 727 and 728 bases of XAC1201 ORF, at amino acid position 244 (Proline). Bacteria were cultivated in NB medium (3.0 g/L beef extract, 5.0 g/L peptone, 15.0 g/L agar) at 28°C for 72 h, supplemented with 35 µg/mL kanamycin when appropriate.

Pathogenicity test *in planta*

Bacterial suspensions of Xac 306 and XAC1201 mutant were adjusted to an $OD_{600\ nm}$ of 0.3, approximately 10^8 CFU/mL, and injected with a needleless syringe onto young and fully expanded leaves of Rangpur lime (*C. limonia* L. Osbeck) and Pera sweet orange (*C. sinensis* L. Osbeck). The cell suspensions were inoculated in three independent leaves, and the test was repeated three times. Disease symptoms were photographed at 3, 5, 7, 9, 12, 15, and 18 days post-inoculation (dpi).

Growth curves

Overnigth cultures of Xac 306 and XAC1201 mutant in NB medium (3.0 g/L beef extract, 5.0 g/L peptone, 15.0 g/L agar) were adjusted to an $OD_{600\ nm}$ of 0.1, using a Spectrophotometer DU-730 (Beckman Coulter). Cultures were incubated in Synergy H1 Hybrid Microplate Reader (BioTek) at 29°C, and absorbance measured every 30 min for 24 h. Two independent experiments were conducted with three replicates each time. The growth curve and the error deviations were plotted using the GraphPad Prism 6 software.

Heterologous expression and purification of the XAC1201 protein

XAC1201 ORF (GenBank™: 1155272) was amplified by PCR with primers forward (5'-ATGAAGTTGGAGGCGTTG-3') and reverse (5'-TTAGTGCAGCAACTCGGC-3') and the PCR product was cloned into pET SUMO expression vector (Invitrogen), according to manufacturer instructions. The expression plasmid was transformed into electrocompetent *E. coli* DH10B cells for amplification and then transformed into One Shot *E. coli* BL21(DE3) pLysS competent cells to express XAC1201 recombinant fusion protein (6xHis-Sumo protein-XAC1201). For expression, one colony of *E. coli* BL21(DE3) pLysS containing the expression vector was inoculated into 15 mL LB medium (10.0 g/L tryptone, 5.0 g/L yeast extract, 10.0 g/L NaCl) supplemented with 35 µg/mL kanamycin and incubated overnight (14-16 h) at 37°C and 250 rpm. The overnight culture (10 mL) was added to 1 L LB medium (10.0 g/L tryptone, 5.0 g/L yeast extract, 10.0 g/L NaCl) supplemented with 35 µg/mL kanamycin and then incubated at 37°C and 250 rpm until reaching an optical density of 0.4 ($OD_{600\ nm}$). At this point, the temperature was lowered to 28°C, and after 10 min at this temperature, the expression was induced by adding IPTG to a final concentration of 0.2 mM. After 4 h, the cells were harvested by centrifugation (CR22GIII, Hitachi) at 1960 *g* for 10 min at 4°C and resuspended in 100 mL lysis buffer (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 5% glycerol, 0.5% Tween 20). The cells were disrupted for 5 min while immersed in an ice-water bath using the microprobe of a Branson's ultrasonic generator (250/450 - Sonifier, Branson) set at 50% maximal energy, using pulses of 10 s followed by 30 s rest, and then centrifuged at 30,000 *g* for 15 min at 4°C (CR22GIII centrifuge, Hitachi). XAC1201 recombinant protein purification was performed by metal affinity chromatography in an Akta Prime Plus FPLC System (GE, Amersham). A cleared lysate (100 mL) was applied to a 5 mL Ni-Sepharose® His-Trap Chelating column (GE, Amersham) equilibrated with buffer A (50 mM Tris-HCl, pH 8.0, 500 mM NaCl, 5% glycerol, 30 mM imidazole). The column was then washed with the same buffer A until the $OD_{280\ nm}$ reached a baseline. The bound proteins were eluted with 60 mL of a continuous gradient with buffer A and buffer B (50 mM Tris-HCl, pH 8.0, 500

mM NaCl, 5% glycerol, 500 mM imidazole). The collected eluted fractions were pooled and digested with SUMO protease in the presence of 1 mM dithiothreitol (DTT) at 4°C for 16 h. After cleavage, the native recombinant protein was separated from the 6xHis tag and Sumo protein (6xHis-Sumo protein) using a gravity column containing 5 mL Ni-Sepharose® High-Performance Media (GE, Healthcare) resin and the elution fraction containing the protein in the native form was concentrated using 10-kDa Amicon filter (Centriprep Concentrators - Amicon). Each step of purification was analyzed by SDS-PAGE. The molecular mass of the XAC1201 protein, predicted by the Expasy program (http://web.expasy.org/compute_pi/), is 29.35 kDa.

Mass spectrometry analysis

The XAC1201 protein band was excised from the SDS-PAGE, and the *in-gel* digestion was performed according to Shevchenko et al. (2006). For this, gel pieces were destained using a solution of 100 mM ammonium bicarbonate/acetonitrile (1:1, v/v) for 4 h before reduction and alkylation using 100 mM dithiothreitol and 50 mM iodoacetamide, respectively. *In-gel* trypsin digestion was carried out for 16 h using a 5-ng/µL digestion solution (10 mM ammonium bicarbonate, 20% acetonitrile in water). Finally, tryptic peptides were extracted from the gel pieces and transferred to a new microtube by using a solution containing 5% formic acid/acetonitrile (1:2, v/v). Before mass spectrometry analyses, peptide samples were dried down in a vacuum centrifuge, resuspended in 20 µL 0.1% formic acid, and separated by liquid chromatography using an EASY-nLC1000 system equipped with a C18 nanocolumn (15 cm, 50 µm ID, 2 µm) and a flow rate of 300 nL/min. Tandem mass spectrometry analysis was carried out in a Q-Exactive system. Mass spectra were acquired in the profile, positive ion mode, and acquisition cycle consisted of one full scan (70,000 FWHM) followed by a maximum of 10 data-dependent scans (17,500 FWHM). Peptide fragmentation was performed by HCD using normalized collision energy (NCE) equal to 35 eV. Confirmation of the XAC1201 protein identity was achieved by searching the experimental MS/MS spectra against the translated *X. citri* subsp *citri* (NCBI taxonomy ID 1203436) genome. The spectral correlation was performed by the Sequest search using the following parameters: 10 ppm and 0.1 Da of precursor and fragment mass tolerance, respectively, oxidation of methionines as a dynamic modification, and carbamidomethylation of cysteines as a static modification. Peptide-spectrum matches were filtered to consider only those with a minimum XCorr equal to 1.5, 2, and 2.5 for singly, doubly, and triply charged ions.

Phosphatase, phosphodiesterase, and pyrophosphohydrolase activity

The method for the phosphatase, phosphodiesterase, and pyrophosphohydrolase assays was adapted from Yakunin et al. (2004), Morar et al. (2015), and Jeon et al. (2016). The reactions were performed for 1 or 2 h at 37°C in 50 mM HEPES, 30 mM KCl buffer, pH 7.0 or 7.5, in the presence of 5 mM $MgCl_2$, 5 mM $MnCl_2$ or 5 mM $NiSO_4$, in a final volume of 0.4 or 1.0 mL. Phosphatase activity was determined in the presence of 5 mM ρ-nitrophenyl phosphate (ρNPP) as substrate. The reaction was interrupted with 1 M NaOH and the released ρ-nitrophenol was measured at 410 nm. Phosphodiesterase assays were performed with 5 mM bis-ρ-nitrophenyl phosphate (bis-ρNPP). The reaction was interrupted by adding 100 µL 50% trichloroacetic acid (TCA) and immediately transferred to ice. The free inorganic phosphate (Pi) was determined according to Heinonen and Lahti (1981), with some modifications. After

the incubation time, the following solutions were added sequentially to each tube: 0.5 mL 0.05 M ammonium molybdate in 2.6 M sulfuric acid, 1.0 mL acetone and 0.5 mL 0.4 M citric acid. After each addition, the tube was vortexed for 15 s, and finally, the absorbance was measured at 355 nm. To test phosphodiesterase activity with natural substrates, 2 mM cyclic adenosine monophosphate (cAMP; Sigma) was added to the reaction mixture with 0.02 U *Saccharomyces cerevisiae* inorganic pyrophosphatase (Sigma). The reaction was stopped with TCA and Pi was measured as described above for phosphodiesterase. To determine whether the XAC1201 protein has pyrophosphohydrolase activity, the protein was incubated with dNTP mix (0.5 mM of each dNTP) in the presence of 5 mM $MgCl_2$, 5 mM $MnCl_2$ or 5 mM $NiSO_4$ and 0.02 U *S. cerevisiae* inorganic pyrophosphatase. The pyrophosphohydrolase activity was also examined by incubation with 0.5 mM of each four dNTP individually (dATP, dTTP, dCTP, or dGTP) in the presence of 5 mM $MgCl_2$ or 5 mM $NiSO_4$ and the presence or absence of 0.5 mM GDP. The reactions were interrupted, and Pi measured as described above. The phosphatase and phosphodiesterase activities using ρNPP and bis-ρNPP as substrates, respectively, were tested in duplicate. In each assay, a control to evaluate the substrate spontaneous hydrolysis was included. One unit (U) of phosphatase activity was defined as the amount of ρ-nitrophenol released (nmol) per hour and per mL. One unit (U) of phosphodiesterase and pyrophosphohydrolase activities were defined as the amount of Pi released (nmol) per hour and per mL.

Circular dichroism (CD) spectroscopy

CD spectra were recorded with a Jasco 810 Spectrophotopolarimeter at 25°C under constant nitrogen flux using a 0.1-cm quartz cuvette. CD spectra of the protein XAC1201 (0.4 mg/mL) were measured in 10 mM Tris-HCl buffer, pH 8.0, after exhaustive dialysis. The CD spectrum of each sample was measured 3 times at a scan rate of 50 nm/min, from which a new buffer spectrum was subtracted. The secondary structure content was estimated by deconvolution of the buffer corrected CD spectra using the package CDPro software. CD signal was expressed in relative Ellipticity [θ]. Data are reported as the mean of triplicate measurement from different enzyme preparations.

Protein modeling

The XAC1201 protein sequence was retrieved from the *Xanthomonas* sp database (https://iant.toulouse.inra.fr//bacteria/annotation/cgi/xansp.cgi) and its three-dimensional (3-D) model was constructed using an iterative implementation of the 'Threading ASSEmbly Refinement' (I-TASSER) program (http://zhanglab.ccmb.med.umich.edu/I-TASSER/). Subsequently, model validation of stereo-chemical quality was carried out using RAMPAGE (http://mordred.bioc.cam.ac.uk/~rapper/rampage.php), ERRAT (http://services.mbi.ucla.edu/ERRAT/), and ProQ (http://www.sbc.su.se/~bjornw/ProQ/) servers. The validation criterion of the model is summarized in **Table S1**. PyMOL (PyMOL Molecular Graphics System, Version 1.2r3pre, Schrödinger, LLC) was used for graphical visualization of all 3-D structures.

In silico analysis

Multiple alignments of the XAC1201 protein with homologous regions of other HD

domain proteins were generated with the ClustalW program (http://www.ch.embnet.org/software/ClustalW.html) and used as the input for the ESPript program, version 3 (http://prodes.toulouse.inra.fr/ESPript/cgi-bin/ESPript.cgi). The ESPript 3 program was also used to visualize the secondary structure of the XAC1201 protein. Protein-protein interactions were predicted by STRING database (version 10.0), with a minimum confidence score of 0.4 using the XAC1201 protein sequence as a query. STRING data sources come from a combination of several computational methods and experimental analysis to predict functional associations between proteins.

Motility assays

The motility assays were performed as described previously (Li and Wang, 2011). Briefly, Xac 306 and the XAC1201 mutant were cultured overnight in NB medium and adjusted to an $OD_{600 \, nm}$ of 0.3 (approximately 10^8 CFU/mL). Two microliters of bacterium culture was spotted onto 15 x 80-mm Petri dishes containing 20 mL NA medium (10 g/L yeast extract, 5 g/L NaCl) with 0.3% (w/v) agar for swimming motility or with 0.7% (w/v) agar for swarming motility. After 72 h at 28°C, the colony diameter was measured using the ImageJ software. The test was repeated three times with three replicates each.

In vitro biofilm formation

The biofilm assays were performed as described by Yan et al. (2012). Briefly, 2-day bacterial cells growing at 29°C were collected from NA plates and re-inoculated into NB medium to obtain a final concentration of 10^8 CFU/mL ($OD_{600 \, nm}$ of 0.3). One milliliter of the cell suspension was added to 20 mL sterilized borosilicate glass tubes. The tubes were kept for 48 h in a humidified chamber at 29°C without shaking. The cultures were then discarded and the tubes were washed two times using tap water. The biofilm formed on the tubes was visualized by staining with 0.1% crystal violet. The stain remaining was dissolved in 1.0 mL of an ethanol-acetone (80:20, v:v) solution and quantified by measuring the optical density at 595 nm using an iMark™ Microplate Absorbance Reader (Bio-Rad Laboratories, Inc., USA). The test was repeated three times with three replicates.

RESULTS

## Pathogenicity profile of the XAC1201 mutant *in planta*

The disease progression was observed from 3 to 18 dpi in two different hosts (Figure 1). The XAC1201 mutant delayed the appearance of symptoms and produced less severe symptoms compared with typical canker disease symptoms caused by the wild-type Xac 306. In Pera sweet orange, a host less susceptible to citrus canker than Rangpur lime, canker symptoms caused by Xac 306 was observed in leaves at 5 dpi, but for the XAC1201 mutant, similar symptoms were observed only at 18 dpi. In Rangpur lime leaves, canker symptoms were observed at 5 dpi for Xac 306 but only at 12 dpi similar symptoms were observed for XAC1201 mutant.
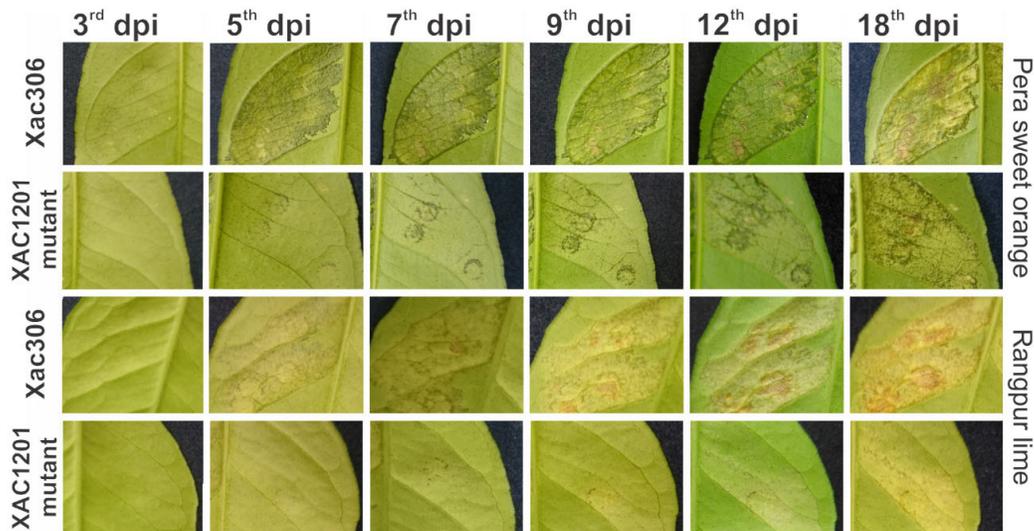
Figure 1. Pathogenicity assay in leaves of two different citrus species. Symptoms observed up to 18 days post-inoculation (dpi) on the abaxial surface of Pera sweet orange (*Citrus sinensis* L. Osbeck) or Rangpur lime (*Citrus limonia* L. Osbeck) leaves infiltrated with bacterial suspensions (10^8 CFU/mL) of either the wild-type Xac 306 or the XAC1201 mutant.

Growth curves

The growth curves of Xac 306 and the XAC1201 mutant revealed similar behavior in a rich environment like NB medium (Figure 2). The bacteria showed similar growth in lag phase (30 min to 2.0 h), then, continue to growth similarly until, approximately, the end of log phase, that lasted until 15 h of growth, in which they started stationary phase until the end of the 24th hour of measurements. However, the Xac1201 mutant did not reach the same absorbance in stationary phase, resulting in a lower growth curve.
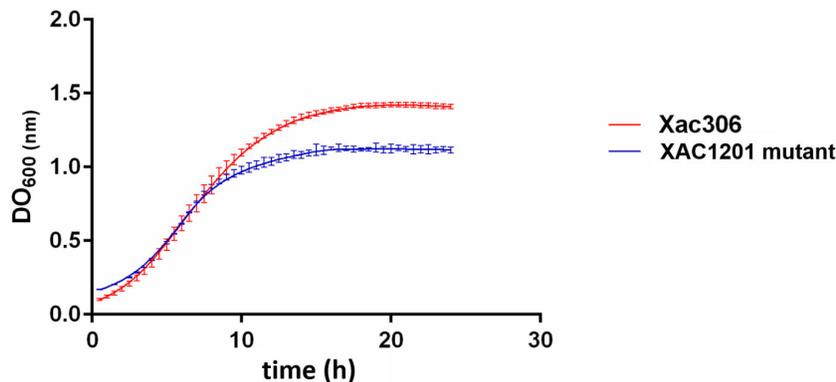
Figure 2. Growth curves of Xac 306 and the Xac1201 mutant in NB medium. Two independent experiments were conducted with three replicates each time. The growth curve and the error deviations were plotted using the GraphPad Prism 6 software.

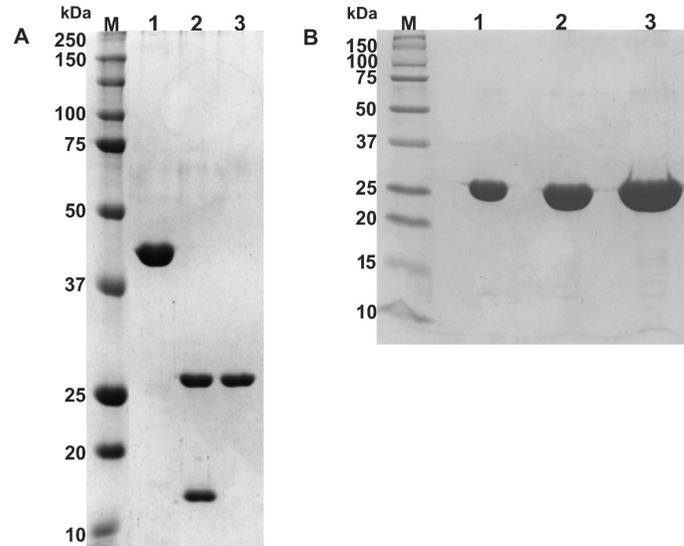## Expression and purification of the XAC1201 recombinant protein

The XAC1201 recombinant fusion protein, having a 6xHis-tag and SUMO protein fused to its N-terminal (6xHis-Sumo protein-XAC1201), has 391 amino acids and a predicted molecular mass of 42.6 kDa. The XAC1201 recombinant protein itself has 272 amino acids and a predicted molecular mass of 29.3 kDa, while the 6xHis-tag and the Sumo protein sequence has 119 amino acids and a predicted molecular mass of 13.3 kDa. Immobilized Metal Affinity Chromatography (IMAC) yielded, in only one step, the XAC1201 fusion protein with high purity, eluted in a single peak in the chromatogram (Figure 3). Cleavage of the XAC1201 fusion protein with SUMO protease renders the XAC1201 recombinant protein without the 6xHis-tag and the Sumo protein. SDS-PAGE analysis showed that the XAC1201 recombinant protein could be eluted from IMAC as a single and highly pure band, with the expected molecular mass of 29.3 kDa (Figure 4).
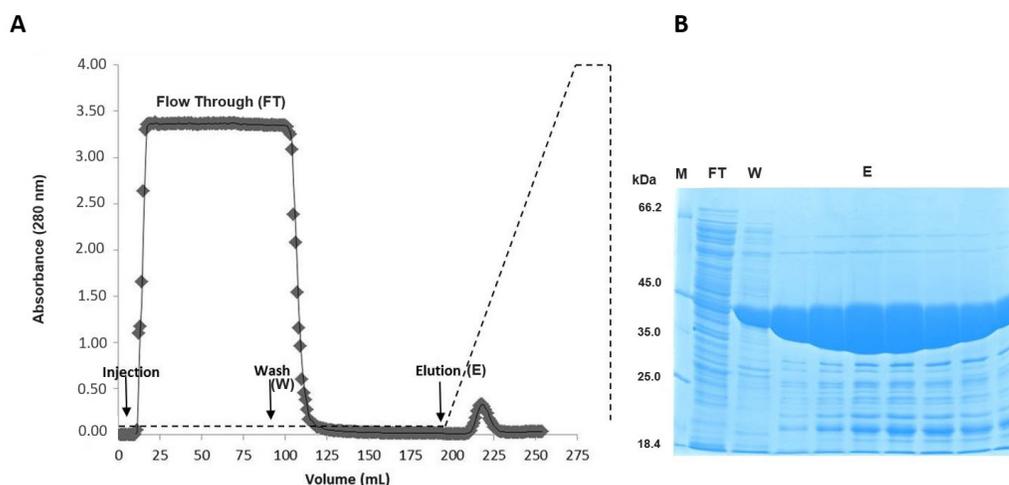
Figure 3. A. Chromatographic profile of *Escherichia coli* lysate containing the XAC1201 recombinant fusion protein in a 5.0 mL Ni^{2+}-Sepharose® column. The loading and washing buffer A were the same (50 mM Tris-HCl, pH 8.0, 30 mM imidazole, 500 mM NaCl, 5% glycerol). The gradient elution (dashed line) was from 100% buffer A to 100% buffer B (50 mM Tris-HCl, pH 8.0, 500 mM imidazole, 500 mM NaCl, 5% glycerol). The flow rate was 2.0 mL/min, and the elution volume was 60 mL. Each fraction contained 2.0 mL eluate. **B.** SDS-PAGE elution profile on a 12% gel. The gel was stained with Coomassie Brilliant Blue R-250. This profile is representative of several others. *Lane M*: molecular marker weight (Unstained Protein Molecular Weight Marker® - Fermentas); *lane FT*: flow-through fraction; *lane W*: washing fraction; *lane E*: elution fractions.

Confirmation of the XAC1201 protein identity

The 29.3-kDa XAC1201 protein band excised from the polyacrylamide gel and digested with trypsin was analyzed by tandem mass spectrometry (MS/MS). Accession XAC1201 was the only hit suggested by the Sequest search engine, identified by 14 unique peptide sequences (**Table S2**). **Figure S1** shows the fragmentation patterns (b- and y-ion series) for the most abundant peptide, MPLLVAQAVLVcDALHAHGGANAAALEAIR.

Figure 4. Purification of the XAC1201 recombinant protein from the recombinant fusion protein 6xHis-Sumo protein-XAC1201. **A.** SDS-PAGE profile of the proteins before and after protease cleavage. *Lane M* = molecular marker weight (Precision Plus Protein Standards kaleidoscope® - Bio-rad); *lane 1* = 5 µg purified XAC1201 recombinant fusion protein before digestion with SUMO protease; *lanes 2* = 5 µg XAC1201 recombinant fusion protein after digestion with SUMO protease; *lanes 3* = 2.5 µg recombinant XAC1201 protein eluted from Ni^{2+}-Sepharose® column after SUMO protease digestion. **B.** SDS-PAGE of purified XAC1201 recombinant protein. *Lane M* = molecular marker weight (Precision Plus Protein Standards kaleidoscope® - Bio-rad); *lane 1* = 18.4 µg protein; *lanes 2* = 36.8 µg protein; *lanes 3* = 73.6 µg protein. Both gels were 12% and were stained with Coomassie Brilliant Blue R-250.

Absence of phosphatase, phosphodiesterase, and pyrophosphohydrolase activities

The XAC1201 protein has a putative conserved HD domain predicted to have metal-dependent phosphohydrolase activity. To verify the presence of such activity, the XAC1201 protein was assayed in the presence of different substrates, cofactors, and incubation conditions. However, no significant activity of phosphatase, phosphodiesterase or pyrophosphohydrolase was detected (Table 1), suggesting that HD domain present in XAC1201 is non-functional.

Table 1. Substrates and cofactors used in enzymatic assays of the XAC1201 protein.

XAC1201 enzymatic activity (U/mL)							
Cofactors	MgCl$_2$	NiSO$_4$	MnCl$_2$	MgCl$_2$	NiSO$_4$	MnCl$_2$	MgCl$_2$ + GDP[c]	NiSO$_4$ + GDP[c]
Substrates	pH 7.0			pH 7.5			pH 7.0	
ρNPP[a]	18.182 ± 0.915	84.091 ± 3.365	-	30.682 ± 1.890	56.818 ± 2.450	-	-	-
bis-ρNPP[a]	7.150 ± 0.063	21.450 ± 1.815	-	100.100 ± 3.505	207.350 ± 8.150	-	-	-
cAMP[b]	9.533 ± 0.826	0.000 ± 0.000	0.000 ± 0.000	9.533 ± 1.092	0.000 ± 0.000	7.150 ± 0.715	-	-
dNTP mix[b]	114.400 ± 10.010	52.433 ± 4.066	224.033 ± 16.773	76.267 ± 7.198	16.683 ± 2.298	90.567 ± 7.973	-	-
dATP[b]	57.200 ± 9.295	109.633 ± 7.6.45	-	-	-	-	130.289 ± 7.355	106.456 ± 3.648
dGTP[b]	11.917 ± 1.092	171.600 ± 28.493	-	-	-	-	34.956 ± 2.614	26.217 ± 2.890
dTTP[b]	30.983 ± 1.488	4.767 ± 0.826	-	-	-	-	9.533 ± 1.092	50.050 ± 2.477
dCTP[b]	76.267 ± 1.488	28.600 ± 2.578	-	-	-	-	98.511 ± 1.397	83.417 ± 6.945

[a]The reactions were performed for 2 h, in a final volume of 1.0 mL, and the absorbance was measured at 410 nm; [b]the reactions were performed for 2 h, in a final volume of 0.4 mL, and absorbance was measured at 355 nm; [c]the reactions were performed for 1 h. The values are reports as means ± SD.

Modeling of the XAC1201 protein structure and sequence alignment reveals a non-functional HD domain

The 3-D molecular model of the XAC1201 protein obtained by the I-TASSER server showed the catalytic site residues of HD domain and its interaction with the predicted cofactors (Figure 5A). According to the model, magnesium (Mg^{2+}) and guanosine diphosphate (GDP) are the main ligands of the XAC1201 protein, with residues H106, H135, N136, V218, and D222 contacting magnesium (**Figure S2A**), and residues H106, N136, G161, H162, A163, A164, E165, Y174, Q198, V199, and V218 contacting GDP (**Figure S2B**). The iron ligand was also identified as a ligand in these analyses, although with a lower score than magnesium (data not shown).
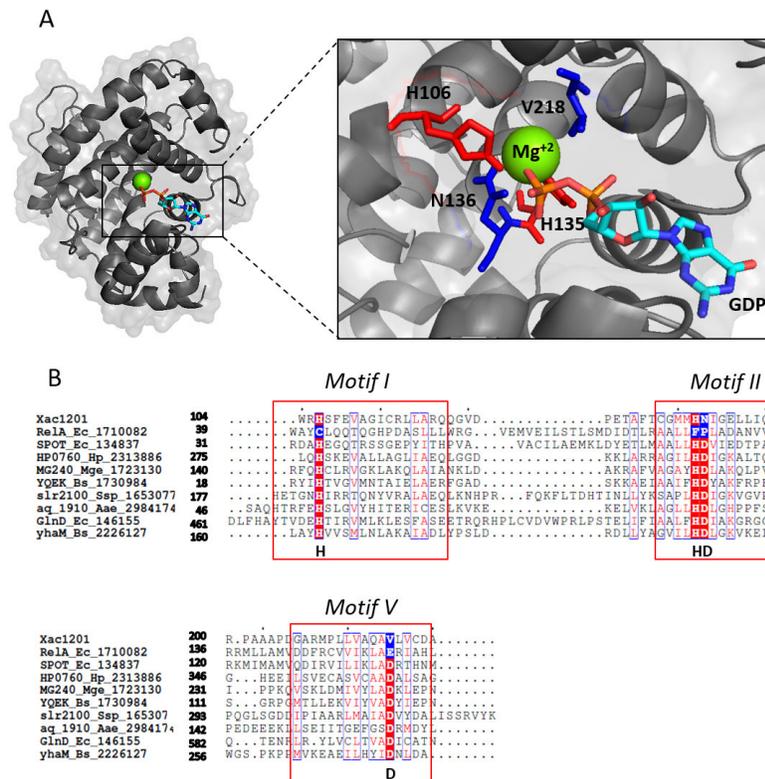
Figure 5. 3-D molecular model of the XAC1201 protein and functionally conserved residues of HD domain. **A.** 3-D model of the XAC1201 protein with magnesium (Mg^{2+}) and guanosine diphosphate (GDP) ligands obtained from the modeling in I-TASSER server. In detail, catalytic site residues are evidenced in stick representation, conserved HD consensus residues are colored in red (H106, H135), and substitutions of the Aspartate residues in the HD domain of XAC1201 are colored in blue (N136, V218). **B.** Multiple sequence alignments of the XAC1201 protein with homologous regions of other HD domain proteins. Sequence regions that correspond to the conserved motifs of the HD superfamily are evidenced in red boxes, and the HD consensus sequence is shown under the alignment. Functionally conserved residues are highlighted in red (color-coding as in A), substitution residues in blue (color-coding as in A), and conserved residues are shown in red font. Multiple alignments were performed using the ESPript 3 server (http://espript.ibcp.fr/ESPript/ESPript/).

      Interestingly, there are two substitutions of the aspartate (D) residues in the catalytic site of the HD domain of XAC1201 (N136, V218), while the histidine (H) residues involved in the catalytic site (H106, H135) are all conserved, despite the fact that no activity was detected in tested conditions. To gain more insights about the non-functional HD domain of the XAC1201 protein, we performed a comparison based on multiple sequence alignments of the XAC1201 protein with homologous regions of the sequences of several organisms, mostly uncharacterized HD-superfamily members that either occur as standalone proteins or associated with other domains (Figure 5B). Among these sequences is the RelA protein from *E. coli* (RelA_Ec_1710082), one of the few members of the non-functional HD domain protein family reported in the literature that possesses substitutions in all conserved catalytic site residues.

## XAC1201 protein structure analysis reveals a major α-helical content

      The CD spectra of the XAC1201 recombinant protein revealed a profile typical for proteins containing significant α-helical content (Figure 6A). Deconvolution of the CD spectra using the CDPro software yielded values of 35-40% of the α-helical content. In addition, the 3-D molecular model of the XAC1201 protein aligned with the crystal structure of two proteins having HD domain [a protein from *Geobacter sulfurreducens* (np_953345.1) (PDB ID 3HC1), with unknown function, and a protein from *Shewanella amazonensis* (Yp_929327.1) (PDB ID 3M1T), with putative phosphohydrolase activity], showed the abundance of α-helices present in those sequences (Figure 6B). There are sixteen α-helices, with no β-pleated sheets. The sequence alignment of the XAC1201 protein with 3HC1 and 3M1T protein sequences, performed by the ESPript 3 server, revealed the presence of the HD domain in the same position in the three protein sequences and the conserved residues within this domain (Figure 6C). However, XAC1201 and 3M1T proteins present the two substitutions of aspartate (D) residues in the catalytic site of the HD domain.

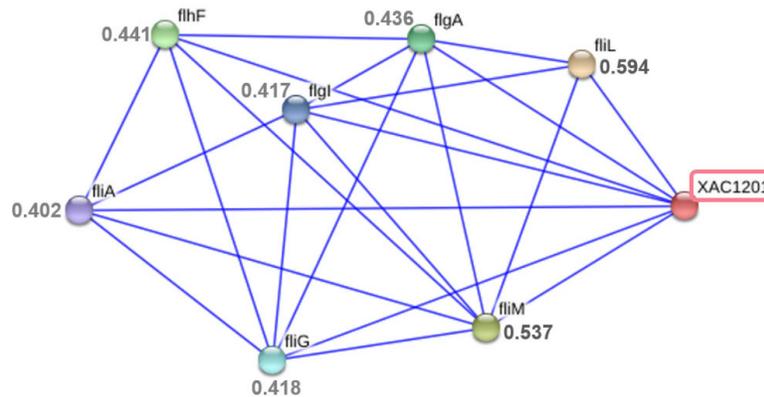## XAC1201 protein interacts with flagella proteins

      To shed light on the XAC1201 protein-protein interactions, STRING analysis (version 10.0) was performed. The result, using a minimum confidence score of 0.4, was a prediction of interaction between XAC1201 and flagellar proteins FliL, FliM, FlgA, FlgL, FlhF, FliA, and FliG (Figure 7). Nonetheless, using a high-confidence score up to 0.5, resulted in a restricted interaction between the XAC1201 protein and FliM and FliL proteins, with 0.537 and 0.594 scores, respectively. To verify if the interaction among the XAC1201 protein and flagella proteins predicted by STRING analysis in fact occurs, the mutant XAC1201 swarming and swimming motility assays were performed. The results revealed that, in fact, there is a significant difference ($P < 0.005$) between the mutant and the wild-type Xac 306 strain in both types of motility (Figure 8), with the XAC1201 mutant showing more motility than the wild-type Xac 306.

Biofilm formation of the XAC1201 mutant

      Biofilm is known to be an important bacterial trait involved in virulence (Sena-Vélez et al., 2015). To verify the relation between the XAC1201 protein and biofilm formation, we
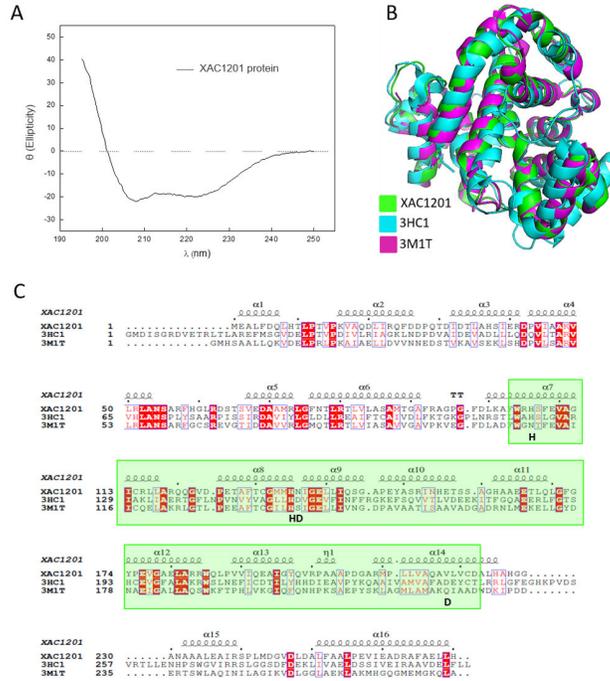
Figure 6. Secondary and tertiary structure of the XAC1201 protein. **A.** Circular dichroism spectra of the XAC1201 protein in the far UV region. **B.** 3-D protein sequence alignment of a molecular model of the XAC1201 protein, in green, with the homologous structures of *Geobacter sulfurreducens* (PDB ID 3HC1), in cyan, and *Shewanella amazonensis* (PDB ID 3M1T) in pink. **C.** Sequence alignment of the XAC1201 protein with the 3HC1 and 3M1T proteins. The HD domain is indicated by the green box, and the HD superfamily consensus sequence is shown under the alignment. Conserved residues are highlighted in red. Multiple alignments were performed using the ESPript 3 server (http://espript.ibcp.fr/ESPript/ESPript/).


Figure 7. STRING analysis of protein-protein interactions with the XAC1201 protein. The network is based on the known and predicted interactions from the STRING database (version 10.0), with a minimum confidence score of 0.4. The blue lines connecting the proteins represent co-occurrence genes (gene families whose occurrence patterns across genomes show similarities).

tested the biofilm formation by mutant XAC1201 and by the wild-type Xac 306 strain in glass tubes, but no significant difference ($P < 0.005$) was observed (Figure 9), suggesting that the XAC1201 protein not stayed involved in biofilm formation.
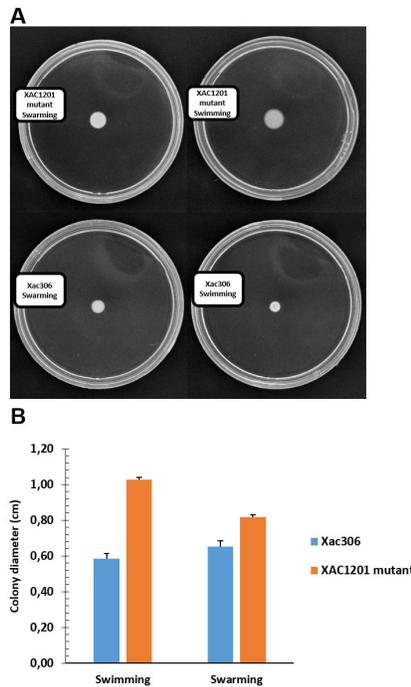
Figure 8. Swimming and swarming motility assays of mutant XAC1201 and wild-type Xac 306. The experiment was repeated three times with three replicates. **A.** Representative photograph of each assay. **B.** The colony diameter was measured with the ImageJ software. Means and standard errors are shown.
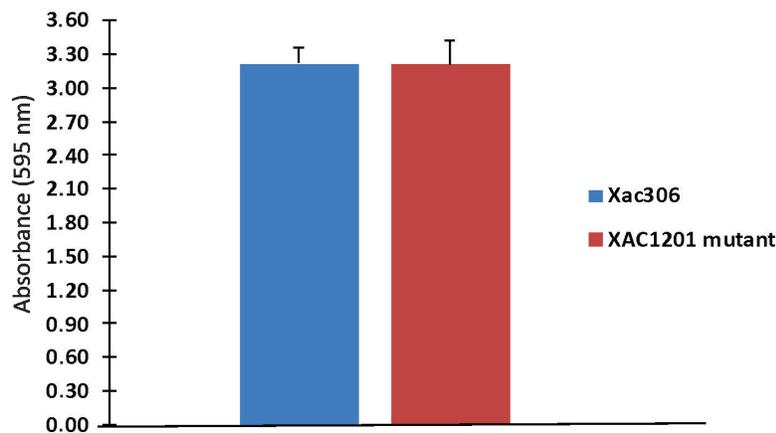
Figure 9. Biofilm formation by XAC1201 mutant and the wild-type *Xanthomonas citri* subsp *citri* strain 306. The experiment was repeated three times with three replicates.

DISCUSSION

We confirmed that disruption of ORF XAC1201 reduces Xac 306 virulence in two different citrus hosts, delaying the onset of typical symptoms, although the *in vitro* growth curves of both bacteria are similar. We also investigated the enzymatic activity, and analyzed *in silico* the structure of the hypothetical protein XAC1201, to understand how it affects the virulence of Xac 306. Analysis using the STRING database predicted that XAC1201 interacts with bacterial flagellar proteins, which are critical for virulence. Therefore, the Xac1201 mutant motility was assayed to test the hypothesis. The I-TASSER analysis allowed us to propose a model for the tridimensional structure of the protein, which revealed modified cofactor binding sites and suggested that XAC1201 has a non-functional HD domain. This was confirmed by enzymatic assays, which, for the tested substrates, revealed no activity previously associated with HD domains.

Several authors (e.g., Ryan et al., 2011; Morar et al., 2015) demonstrated that proteins lacking functional HD domains are often related to regulation of proteins involved in virulence. The analysis in the STRING database suggested that the XAC1201 protein interacts with the flagellar proteins FliL, FliM, FlgA, FlgL, FlhF, FliA, and FliG (Figure 7), all of which are involved in bacterial flagella structure and function. Two proteins exhibited confidence scores higher than 0.5 for this interaction: FliL and FliM. FliL is a flagellar protein part of flagellar regulon and required for flagellar function (Wolfe and Visick, 2008; Moreira et al., 2015), whereas FliM is a flagellar motor switch protein, one of three proteins (FliG, FliN, FliM) that forms the rotor-mounted switch complex (C ring), located at the base of the basal body (Moreira et al., 2015).

Flagellum-mediated motility is crucial for initial cell attachment to host and is involved in mature biofilm formation (O'Toole et al., 2000). In *X. citri* subsp *citri*, the flagellum is known to be essential for bacterial growth and development of citrus canker (Malamud et al., 2011). Flagellin, the protein subunit of flagellum, is a known PAMP (pathogen-associated molecular pattern) in many plant species and is highly perceived by the host-specific chemo-perception system, resulting in PAMP-triggered immunity (PTI) (Felix et al., 1999; Chisholm et al., 2006), the first level of plant response induced by pathogen recognition. We hypothesize that the strong motility of the XAC1201 mutant, shown in Figure 8, leads to a great amount of flagellin being released into plant cells and then enhancing a strong PTI response; however, bacterial effectors can overcome host defense barriers, resulting in disease. This relation might explain why the XAC1201 mutant, previously obtained by aleatory Tn-5 transposon insertion (Laia et al., 2009), displayed reduced and delayed symptoms in different hosts when compared to the wild-type Xac 306.

We generated a model for the XAC1201 protein using the I-TASSER server. This model demonstrates that XAC1201 does not carry a whole HD domain. The HD domains have three conserved motifs (I, II, and V) in the entire superfamily that contain the catalytic core. The catalytic core composition includes four highly conserved residues, two histidine (H) and two aspartate (D) residues (H…HD…D) and requires the presence of a divalent cation to perform its function (Aravind and Koonin, 1998; Zimmerman et al., 2008). The sequence alignment comparison of the XAC1201 protein with various HD domain proteins revealed that the XAC1201 HD domain catalytic site residues are not conserved. The exchange of Asp (D) for Asn (N136) and Asp for Val (V218) residues in the catalytic site (Figure 5), important for metal coordination, led to a possible electrostatic change that compromises the

metal binding interaction in the catalytic site residues, affecting its activity. These results are in agreement with our enzymatic assays. Although we were able to clone, express in *E. coli*, purify to homogeneity, and conduct several enzymatic assays on substrates as ρNPP, bis-ρNPP, cAMP, and dNTPs (dATP, dGTP, dCTP, and dTTP), all of which are known to act as substrates for phosphohydrolases, in the presence of Mg^{2+}, Mn^{2+}, and Ni^{2+}, as metal ligands, no significant enzymatic activity was detected.

Interestingly, a few members of the HD superfamily protein are reported to have an inactive catalytic site, due to alterations in H and/or D residues within the catalytic site. For example, the guanosine 5'-diphosphate 3'-diphosphate (ppGpp) produced by the RelA protein from *E. coli* that helps to regulate several factors in response to nutritional stress, has lost its hydrolase activity with substitutions in the His, Asp, and Glu residues (Potrykus and Cashel, 2008). A variant of HD-GYP domain (YN-GYP), from *Pseudomonas aeruginosa*, involved in bacterium motility, is enzymatically inactive, lacking cyclic-di-GMP phosphodiesterase activity (Ryan et al., 2011). Recently, a bacterial effector with unknown function from the Type 4B Secretion System (T4BSS) of *Legionella pneumophila* showed no phosphohydrolase enzymatic activity due to the change of the His-Asp catalytic residues typical of the HD domain, to Ser and Val (Morar et al., 2015). Therefore, according to our data, we propose that the XAC1201 protein should be considered a new member of the particular group of HD domains carrying an inactive catalytic site.

Our study showed that the HD domain present on the XAC1201 protein does not possess phosphohydrolase activity, although mutant carrying disruption on the XAC1201 gene decreased bacterium virulence in the citrus host. The suggestion is that the XAC1201 protein is related to the flagellar proteins of Xac 306, which has been shown to be necessary for bacterium virulence. This does not exclude the involvement of the XAC1201 protein in other virulence pathways.

ACKNOWLEDGMENTS

REFERENCES

Aravind L and Koonin EV (1998). The HD domain defines a new superfamily of metal-dependent phosphohydrolases. *Trends Biochem. Sci.* 23: 469-472. https://doi.org/10.1016/S0968-0004(98)01293-6

Brunings AM and Gabriel DW (2003). *Xanthomonas citri*: breaking the surface. *Mol. Plant Pathol.* 4: 141-157. https://doi.org/10.1046/j.1364-3703.2003.00163.x

Chisholm ST, Coaker G, Day B and Staskawicz BJ (2006). Host-microbe interactions: shaping the evolution of the plant immune response. *Cell* 124: 803-814. https://doi.org/10.1016/j.cell.2006.02.008

da Silva ACR, Ferro JA, Reinach FC, Farah CS, et al. (2002). Comparison of the genomes of two *Xanthomonas* pathogens with differing host specificities. *Nature* 417: 459-463. https://doi.org/10.1038/417459a

Das AK (2003). Citrus Canker - A review. *J. Appl. Hortic.* 5: 52-60.

Felix G, Duran JD, Volko S and Boller T (1999). Plants have a sensitive perception system for the most conserved domain of bacterial flagellin. *Plant J.* 18: 265-276. https://doi.org/10.1046/j.1365-313X.1999.00265.x

Graham JH, Gottwald TR, Cubero J and Achor DS (2004). *Xanthomonas axonopodis* pv. *citri*: factors affecting successful eradication of citrus canker. *Mol. Plant Pathol.* 5: 1-15. https://doi.org/10.1046/j.1364-3703.2004.00197.x

Heinonen JK and Lahti RJ (1981). A new and convenient colorimetric determination of inorganic orthophosphate and its application to the assay of inorganic pyrophosphatase. *Anal. Biochem.* 113: 313-317. https://doi.org/10.1016/0003-2697(81)90082-8

Jeon YJ, Park SC, Song WS, Kim O-H, et al. (2016). Structural and biochemical characterization of bacterial YpgQ protein reveals a metal-dependent nucleotide pyrophosphohydrolase. *J. Struct. Biol.* 195: 113-122. https://doi.org/10.1016/j.jsb.2016.04.002

Laia ML, Moreira LM, Dezajacomo J, Brigati JB, et al. (2009). New genes of *Xanthomonas citri* subsp. *citri* involved in pathogenesis and adaptation revealed by a transposon-based mutant library. *BMC Microbiol.* 9: 12. https://doi.org/10.1186/1471-2180-9-12

Li J and Wang N (2011). The wxacO gene of *Xanthomonas citri* ssp. *citri* encodes a protein with a role in lipopolysaccharide biosynthesis, biofilm formation, stress tolerance and virulence. *Mol. Plant Pathol.* 12: 381-396. https://doi.org/10.1111/j.1364-3703.2010.00681.x

Malamud F, Torres PS, Roeschlin R, Rigano LA, et al. (2011). The *Xanthomonas axonopodis* pv. *citri* flagellum is required for mature biofilm and canker development. *Microbiology* 157: 819-829. https://doi.org/10.1099/mic.0.044255-0

Morar M, Evdokimova E, Chang C, Ensminger AW, et al. (2015). Crystal structure of the *Legionella pneumophila* Lem10 effector reveals a new member of the HD protein superfamily. *Proteins* 83: 2319-2325. https://doi.org/10.1002/prot.24933

Moreira LM, Facincani AP, Ferreira CB, Ferreira RM, et al. (2015). Chemotactic signal transduction and phosphate metabolism as adaptive strategies during citrus canker induction by *Xanthomonas citri. Funct. Integr. Genomics* 15: 197-210. https://doi.org/10.1007/s10142-014-0414-z

O'Toole G, Kaplan HB and Kolter R (2000). Biofilm formation as microbial development. *Annu. Rev. Microbiol.* 54: 49-79. https://doi.org/10.1146/annurev.micro.54.1.49

Potrykus K and Cashel M (2008). (p)ppGpp: still magical? *Annu. Rev. Microbiol.* 62: 35-51. https://doi.org/10.1146/annurev.micro.62.081307.162903

Ryan RP, Vorhölter F-J, Potnis N, Jones JB, et al. (2011). Pathogenomics of *Xanthomonas*: understanding bacterium-plant interactions. *Nat. Rev. Microbiol.* 9: 344-355. https://doi.org/10.1038/nrmicro2558

Sena-Vélez M, Redondo C, Gell I, Ferragud E, et al. (2015). Biofilm formation and motility of *Xanthomonas* strains with different citrus host range. *Plant Pathol.* 64: 767-775. https://doi.org/10.1111/ppa.12311

Shevchenko A, Tomas H, Havlis J, Olsen JV, et al. (2006). *In-gel* digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* 1: 2856-2860. https://doi.org/10.1038/nprot.2006.468

Wolfe AJ and Visick KL (2008). Get the message out: cyclic-Di-GMP regulates multiple levels of flagellum-based motility. *J. Bacteriol.* 190: 463-475. https://doi.org/10.1128/JB.01418-07

Yakunin AF, Proudfoot M, Kuznetsova E, Savchenko A, et al. (2004). The HD domain of the *Escherichia coli* tRNA nucleotidyltransferase has 2',3'-cyclic phosphodiesterase, 2'-nucleotidase, and phosphatase activities. *J. Biol. Chem.* 279: 36819-36827. https://doi.org/10.1074/jbc.M405120200

Yan Q, Hu X and Wang N (2012). The novel virulence-related gene nlxA in the lipopolysaccharide cluster of *Xanthomonas citri* ssp. *citri* is involved in the production of lipopolysaccharide and extracellular polysaccharide, motility, biofilm formation and stress resistance. *Mol. Plant Pathol.* 13: 923-934. https://doi.org/10.1111/j.1364-3703.2012.00800.x

Zimmerman MD, Proudfoot M, Yakunin A and Minor W (2008). Structural insight into the mechanism of substrate specificity and catalytic activity of an HD-domain phosphohydrolase: the 5'-deoxyribonucleotidase YfbR from *Escherichia coli. J. Mol. Biol.* 378: 215-226. https://doi.org/10.1016/j.jmb.2008.02.036

Supplementary material

Table S1. Quality of the models predicted by I-TASSER server.

Table S2. Peptide sequences identified for the protein entry XAC1201.

Figure S1. Spectral fragmentation pattern for the peptide MPLLVAQAVLVcDALHAHGGANAAALEAIR. **A.** b-ion series; **B.** y-ion series.

Figure S2. Residues involved in the catalytic site of the XAC1201 protein. **A.** Residues involved in the binding of magnesium (Mg^{2+}) and **B.** in the binding of guanosine diphosphate (GDP). The binding site residues are in stick representation in blue color.