Biochimica et Biophysica Acta 1834 (2013) 697–707

Contents lists available at SciVerse ScienceDirect

Biochimica et Biophysica Acta

journal homepage: www.elsevier.com/locate/bbapap

Small-angle X-ray scattering and *in silico* modeling approaches for the accurate functional annotation of an LysR-type transcriptional regulator

M.A.S. Toledo [a,1], C.A. Santos [a,1], J.S. Mendes [a,1], A.C. Pelloso [a], L.L. Beloti [a], A. Crucello [a], M.T.P. Favaro [a], A.S. Santiago [a], D.R.S. Schneider [a], A.M. Saraiva [a,b], D.R. Stach-Machado [c], A.A. Souza [d], D.B.B. Trivella [e], R. Aparicio [e], L. Tasic [f], A.R. Azzoni [a,g], A.P. Souza [a,h,*]

[a] *Centro de Biologia Molecular e Engenharia Genética, Universidade Estadual de Campinas, Campinas, SP, Brazil*
[b] *Instituto Nacional de Metrologia, Qualidade e Tecnologia, Rio de Janeiro, Brazil*
[c] *Departamento de Imunologia — Instituto de Biologia — Univerisade Estadual de Campinas, Campinas, SP, Brazil*
[d] *Centro APTA Citros Sylvio Moreira/IAC, Rodovia Anhanguera Km 158, Cordeirópolis, SP, Brazil*
[e] *Laboratório de Biologia Estrutural e Cristalografia, Instituto de Química, Universidade Estadual de Campinas, Campinas, SP, Brazil*
[f] *Laboratório de Química Biológica, Instituto de Química, Universidade Estadual de Campinas, Campinas, SP, Brazil*
[g] *Departamento de Engenharia Química, Escola Politécnica, Universidade de São Paulo, São Paulo, SP, Brazil*
[h] *Instituto de Biologia, Universidade Estadual de Campinas, Departamento de Biologia Vegetal, Campinas, SP, Brazil*

ARTICLE INFO

Article history:
Received 23 June 2012
Received in revised form 24 December 2012
Accepted 26 December 2012
Available online 5 January 2013

Keywords:
LysR-like transcriptional regulator (LTTR)
Small-angle X-ray scattering (SAXS)
Benzoate
Muconate
Biofilm
Citrus variegated chlorosis (CVC)

ABSTRACT

Xylella fastidiosa is a xylem-limited, Gram-negative phytopathogen responsible for economically relevant crop diseases. Its genome was thus sequenced in an effort to characterize and understand its metabolism and pathogenic mechanisms. However, the assignment of the proper functions to the identified *open reading frames* (ORFs) of this pathogen was impaired due to a lack of sequence similarity in the databases. In the present work, we used small-angle X-ray scattering and *in silico* modeling approaches to characterize and assign a function to a predicted LysR-type transcriptional regulator in the *X. fastidiosa* (XfLysRL) genome. XfLysRL was predicted to be a homologue of BenM, which is a transcriptional regulator involved in the degradation pathway of aromatic compounds. Further functional assays confirmed the structural prediction because we observed that XfLysRL interacts with benzoate and *cis*,*cis*-muconic acid (also known as 2E,4E-hexa-2,4-dienedioic acid; hereafter named muconate), both of which are co-factors of BenM. In addition, we showed that the XfLysRL protein is differentially expressed during the different stages of *X. fastidiosa* biofilm formation and planktonic cell growth, which indicates that its expression responds to a cellular signal that is likely related to the aromatic compound degradation pathway. The assignment of the proper function to a protein is a key step toward understanding the cellular metabolic pathways and pathogenic mechanisms. In the context of *X. fastidiosa*, the characterization of the predicted ORFs may lead to a better understanding of the cellular pathways that are linked to its bacterial pathogenicity.

1. Introduction

Xylella fastidiosa is a xylem-limited, Gram-negative bacterium and the causal agent of many important crop diseases, including citrus variegated chlorosis (CVC), "phony peach" in peaches, and Pierce's disease in grapes [1–4]. Because of the economic impact of these diseases, the *X. fastidiosa* genome was sequenced [5], which allowed the identification of all of the basic genetic profiles for the survival of this bacterium, including the genes involved in metabolism, the synthesis of essential molecules (*e.g.*, amino acids, nucleotides and lipids), and mechanisms of transcription, translation and repair. However, because putative functions could be assigned to only 47% of the 2904 predicted coding regions, the functions of many *open reading frames* (ORFs) remain unknown [5–8]. Since the sequencing of the *X. fastidiosa* genome, more sophisticated experiments have been performed that have led to a better characterization and understanding of this phytopathogen and its pathogenic mechanisms [9–11]. Currently, the hypothesis for the pathogenicity of the bacterium involves the occlusion of the xylem vessels of the infected plants either through the secretion of exopolysaccharides to form biofilm structures or by mechanical occlusion, which leads to hydric stress and the appearance of disease symptoms, such as leaf marginal necrosis, leaf abscission, delayed growth, and reduced plant vigor and fruit development [12–15]. In addition, it has been reported that the gene

* Corresponding author at: Depto. de Biologia Vegetal, Universidade Estadual de Campinas, Caixa Postal 6010, CEP 13083-875, Campinas, São Paulo, Brazil. Tel.: +55 19 3521 1132.
E-mail address: anete@unicamp.br (A.P. Souza).
[1] These authors contributed equally to this work.

1570-9639/$ – see front matter © 2012 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.bbapap.2012.12.017

expression profile of planktonic *X. fastidiosa* cells is different from that of biofilm-forming cells. The differentially regulated ORFs are likely to be involved in the cell's ability to adhere to xylem vessels and form the biofilm structure that leads to the disease onset. Therefore, investigations of these differentially expressed ORFs have been performed to understand the biofilm formation process and thus, the pathogenicity of *X. fastidiosa* [16,17].

In this study, we characterized the protein product of Xf1448 ORF, which is predicted to be a LysR-like transcriptional regulator (LTTR; http://www.xylella.lncc.br/xf-prod-bin/annotation/final/annotation.cgi?id=&gene=XF1448). Because this ORF has little amino acid sequence similarity with other LysR-type proteins, the identification of its role in the cellular metabolism was previously impaired. The LTTR family is a well-characterized group of transcriptional regulators that is ubiquitous among bacteria, has a high degree of conservation, and has functional orthologues in *archaea* and eukaryotic organisms [18–20]. These transcriptional regulators usually control their own expression, which may occur simultaneously with the regulation of a divergently transcribed target ORF or operon [21]. Based on sequence alignments, it has been shown that LTTRs have a helix-turn-helix (HTH) DNA-binding domain located 20–90 amino acids from the N-terminus, regardless of whether the transcriptional regulator acts as an activator or repressor [21]. Generally, this DNA-binding domain interacts with the promoter region of the targeted ORF or operon through a pseudo-palindromic sequence (T-N_{11}-A) that is called the LTTR box [22]. However, the LTTR box may vary in both base pair composition and length [21]. The C-terminus (amino acids 95–210) of a typical LTTR contains the co-inducer binding cleft, and the interaction between the co-inducer and LTTR modulates the activating or repressing function of the protein [23–25]. The molecules that act as co-inducers are generally by-products or metabolic intermediates of the pathways that are regulated by the LTTRs [26,27]. Considering the abundance of LTTRs within the genomes of evolutionarily distant bacteria and other organisms, LTTRs have developed regulatory functions in many different pathways, including metabolism, cell division, quorum sensing, virulence, motility, nitrogen fixation, oxidative stress response, toxin production, attachment and secretion [9,28–36]. Moreover, this class of tranSmall-angle X-ray scattering and *in silico* modeling approaches for the accurate functional annotation of an LysR-type transcriptional regulatorscription factors may act in association with other regulatory pathways that induce more complex regulatory networks inside the cell. One type of interaction is the association of the LTTR with the globally expressed histone-like nucleoid-associated regulator protein (H-NS) [37].

In the present work, we identified the structure and function of the predicted LysR-type ORF Xf1448 from *X. fastidiosa* (denoted XfLysRL for LysR-like from *X. fastidiosa* in the present work). Although the XfLysRL ORF is described as a LTTR, it does not exhibit sufficient sequence similarity with well-characterized LTTRs (functionally and/or structurally), which had previously impaired a more accurate annotation of its function. However, it does contain a conserved LTTR substrate binding domain that is involved in the catabolism of aromatic compounds. Using small-angle X-ray scattering (SAXS) and *in silico* modeling, we were able to obtain a reliable protein structure envelope, which allowed us to use structural conservation to identify putative homologues with known functions. Further functional assays confirmed the role of XfLysRL in the cell, which allowed the identification of new information and knowledge of the *X. fastidiosa* metabolism.

2. Materials and methods

2.1. Cloning, expression and purification of recombinant XfLysRL

The gene corresponding with XfLysRL was amplified by PCR using genomic DNA from *X. fastidiosa* strain *9a5c* as the template. Specific primers for the target sequence were designed such that the fragments produced could be cloned into the pET28a vector with *Nde*I and *Xho*I (forward primer: 5′ — TAATCATATGCACGACGCCGCCAGT — 3′; reverse primer: 5′ — ATCTCGAGTCACCTTGCACCAGCAC — 3′). The PCR product was cloned into pET28a and chemically transformed into competent *Escherichia coli* DH5-α cells. The positive clones were sequenced to verify the identity and presence of the cloned fragment. The competent *E. coli* BL21(DE3) strain was used for the recombinant protein expression.

E. coli BL21(DE3) cells containing the gene that encodes XfLysRL were inoculated in 3 mL of LB medium containing 30 μg/mL kanamycin. The cells were then grown at 37 °C and 300 rpm overnight. The cultures were transferred into 1 L of LB medium containing the same concentration of antibiotic and grown to an A_{600} of 0.6–0.8. The overexpression of the plasmid was induced through the addition of IPTG to a final concentration of 0.4 mM; the culture was then incubated at 25 °C and 200 rpm for 18 h. After the incubation period, the cells were centrifuged at 5000 × g for 15 min (4 °C), and the pellet was resuspended in buffer A (50 mM Tris–HCl and 300 mM NaCl pH 7.5) supplemented with 1 mg/mL lysozyme and 1 mM phenylmethanesulfonyl fluoride (PMSF; Sigma Chemical, St. Louis, MO, USA). The cells were maintained under agitation conditions for 30 min (4 °C), sonicated and then centrifuged at 15,000 × g for 45 min (4 °C) to obtain the protein extract. The XfLysRL protein was then purified through a single affinity chromatography step using a Ni-NTA resin (Qiagen, Hilden, Germany) equilibrated with buffer A. The purified protein was eluted using an imidazole gradient (20, 50, 75, 100, 200 and 500 mM) in buffer A. The protein yield and purity were analyzed by SDS-PAGE.

2.2. Size-exclusion chromatography

Size-exclusion chromatography was performed using a Superdex 75 10/300 GL column (GE Healthcare, USA). The column was equilibrated with 50 mM Tris–HCl pH 7.5 and 300 mM NaCl. The samples (250 μL) were injected at a flow rate of 0.6 mL/min. High molecular weight (HMW) and low molecular weight (LMW) gel filtration calibration kits (GE Healthcare) were used as the calibration standards, and the results were analyzed according to the instructions detailed in the calibration kit manuals. Conalbumin (75 kDa), ovalbumin (43 kDa), carbonic anhydrase (29 kDa), ribonuclease A (13.7 kDa), aprotinin (6.5 kDa) and blue dextran 2000 (2000 kDa) were used as the standards.

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

The XfLysRL samples at a concentration of 1.5 mg/mL in 50 mM Tris–HCl pH 7.5, 300 mM NaCl and 500 mM imidazole buffer were subjected to SAXS data collection. The samples were centrifuged at 13,000 × g and 4 °C prior to the measurements. The individual samples were carefully loaded into cells composed of two thin parallel mica windows and maintained at 25 °C throughout the measurements. The SAXS data collection was performed at the D02A-SAXS2 beamline of the Brazilian National Synchrotron Light Laboratory (LNLS, Campinas) [38] using a bi-dimensional MAR CCD 345 detector and a monochromatic X-ray beam at a wavelength of 1.488 Å. The sample-to-detector distance was adjusted to 1682.88 mm, which covers a momentum transfer interval of $0.0076 < q < 0.1928$ Å^{-1}, where $q = [(4\pi)/(\lambda)] \sin\theta$ and 2θ is the scattering angle.

The protein samples and buffer solution (used as a blank during the data analysis) were exposed to the X-ray for 10 min. Six successive frames were recorded for the protein sample. A data reduction was performed using FIT2D [39], which involved the collection of the radial integration of the images, normalization to the intensity of the transmitted beam and subtraction of the buffer scattering. To prevent artifacts caused by protein degradation during the data