FEBS Letters 587 (2013) 339–344

FEBS Letters

journal homepage: www.FEBSLetters.org

The crystal structure of the cysteine protease Xylellain from *Xylella fastidiosa* reveals an intriguing activation mechanism

Ney Ribeiro Leite [a,1,2], Aline Regis Faro [a,1,3], Maria Amélia Oliva Dotta [a], Livia Maria Faim [a], Andreia Gianotti [b], Flavio Henrique Silva [b], Glaucius Oliva [a], Otavio Henrique Thiemann [a,*]

[a] *Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, SP 13566-590, Brazil*
[b] *Department of Genetics and Evolution, Federal University of São Carlos, São Carlos, SP 13565-905, Brazil*

ARTICLE INFO

Article history:
Received 11 September 2012
Revised 21 December 2012
Accepted 3 January 2013
Available online 17 January 2013

Edited by Stuart Ferguson

Keywords:
Cysteine protease
Xylellain
Crystal structure
Xylella fastidiosa

ABSTRACT

***Xylella fastidiosa* is responsible for a wide range of economically important plant diseases. We report here the crystal structure and kinetic data of Xylellain, the first cysteine protease characterized from the genome of the pathogenic *X. fastidiosa* strain 9a5c. Xylellain has a papain-family fold, and part of the N-terminal sequence blocks the enzyme active site, thereby mediating protein activity. One novel feature identified in the structure is the presence of a ribonucleotide bound outside the active site. We show that this ribonucleotide plays an important regulatory role in Xylellain enzyme kinetics, possibly functioning as a physiological mediator.**

Structured summary of protein interactions:
Xylellain and **Xylellain** bind by X-ray crystallography (View interaction)

1. Introduction

Xylella fastidiosa is a gram-negative bacteria that colonizes exclusively the plant xylem vessels decreasing xylem flow compromising the plant development and are difficult (*fastidious*) to culture by standard bacteriological procedures [1–3]. Several economically important crops are affected by *X. fastidiosa* strains [4–8] that are transmitted by Homoptera (leafhoppers) insects [9,10]. In Brazil alone, citrus variegated chlorosis or CVC cause approximate annual losses to the citrus industry of US $100 million [11] whose control is limited to crop management, infected branches pruning and plant eradication. The use of healthy seedlings and vector control are at present the only preventive measures available. The availability in 2000 of the *X. fastidiosa* clone 9a5c genome sequence [12] lead to a significant advance in the understanding of the bacterium metabolism [2,3]. From *X. fastidiosa* genome analysis a single cysteine protease was identified, Xylellain, and characterized as a cathepsin B like protease [13]. In this study, we solved the crystal structure of Xylellain, to a resolution of 1.65 Å. This structure revealed that the first 39 residues of the enzyme N-terminal sequence are attached to the active site and represent a regulatory pro-region. Surprisingly a ribonuclotide, UDP, is present in a hinge region and may play an important role in stabilizing the N-terminal pro-region, representing a potential enzyme activity modulator, sensing the bacteria physiological state.

2. Materials and methods

2.1. Recombinant Xylellain characterization

The expression and purification of the recombinant Xylellain protein was performed essentially as previously described [13] modified by the induction temperature reduced to 20 °C for 16 h and isopropyl-b-D-thiogalactopyranoside (IPTG) added to a final concentration of 0.1 mM selenomethionine incorporation for single-wavelength anomalous dispersion (SAD) experiments was performed by inhibition of methionine biosynthesis in M9 minimal medium (Sigma) supplemented by selenomethionine (Sigma) [14].

Point mutations R23A, F25A and R23A/F25A, were generated using the Gene Tailor™ Site-Directed Mutagenesis kit (Invitrogen), according to the manufacturer's instructions and the oligonucleotides: R23A, 5′-ATAGCTGATATTGCTGACTTTTCATACACC-3′; F25A, 5′-GATATTCGTGACGCTTCATACACCCCAGAG-3′; R23A/F25A, 5′-ATAGCTGATATTGCTGACGCTTCATACACCCC-3′ and their complementary primer: R23A and R23A/F25A, 5′-ATAGGGTCTATATCGAC-

* Corresponding author. Fax: +55 1633739881.
E-mail address: thiemann@ifsc.usp.br (O.H. Thiemann).
[1] Both authors contributed equally to this work.
[2] Present address: Cristália Produtos Químicos Farmacêuticos Ltda. Rodovia Itapira-Lindóia, km 14, Itapira, SP, Brazil.
[3] Present address: Institut de Biologie Structurale Jean-Pierre Ebel, Université Joseph Fourier, 41 rue Jules Horowitz, F-38027 Grenoble, France.

0014-5793/$36.00
http://dx.doi.org/10.1016/j.febslet.2013.01.009

TATAA-3′, F25A, 5′-TCTATATCGACTATAAGCACTG-3′. Mutant Xylellain proteins were expressed and purified as described.

Enzyme kinetic experiments were performed as described [13] in a WALLAC 1420 fluorimeter using carbobenzoxy-Phe-Arg-7-amido-4-methylcoumarin (Z-RF-MCA, Sigma) as substrate. The amount of active enzyme was determined by titration with E64 (N-[N-(L-3-trans-carboxyoxirane-2-carbonyl)-L-leucyl]-agmatine) inhibitor [15].

2.2. Crystallization and data collection

Suitable crystals were obtained at 18 °C using the hanging-drop vapor diffusion technique in 60 mM sodium citrate (pH 5.6), 134 mM ammonium acetate and 20–22% (w/v) PEG 4000 and 7 mg/ml of Xylellain. The crystals were cryoprotected in the crystallization solution containing 15% (v/v) ethylene glycol and frozen to 100 K. A native data set was collected to 1.65 Å resolution, on a MAR345dtb image-plate detector using Cu Kα radiation generated by a Rigaku Ultra-X 18 rotating-anode operating at 90 mA and 50 kV and focused using Osmic mirrors. Selenomethionine crystals were obtained by the hanging-drop vapor diffusion method in 100 mM MES (pH 5.6), 18% PEG 8000 and 6 mg/ml of Xylellain. Selenomethionine data sets were collected at the National Synchrotron Light Source, Brookhaven National Laboratory, Beamline X12B. Xylellain diffraction data for the Sel-Met crystal were indexed, integrated and reduced using HKL2000 package [16] while for the native crystal, the data processing was carried out using MOSFLM [17]. Intensities were scaled with SCALA from the CCP4 suite [18]. Data collection and processing statistics are provided in Table 1. Both native and selenomethionine crystals belong to space group P1 with similar cell dimensions and Matthew's coefficient [19] calculated to 2.16 $Å^3$/Da with a solvent content of 43.02%.

2.3. Structure solution and refinement

Searches on the PDB database identified *Cathepsin S* and *F* [20,21] and *Cruzain* [22] as Xylellain homologues with 30% sequence identity. Molecular replacement attempts to solve the Xylellain structure with AmoRe [23], MOLREP [24] and PHASER [25] were not successful justifying the use of selenomethionine isomorphous replacement methodology.

The selenomethionine Xylellain structure was solved by SAD method at the Se-K edge (λ = 0.9792). Eight selenium atoms per cell unit were found using SHELX package [26], initial phases were obtained with SOLVE [27] and improved by solvent flattening and density modification with 42% solvent content, followed by auto building routine with RESOLVE [28]. Residues of the initial model were checked according to their electronic density with COOT [29]. Initial phases were extended with sequential rounds of solvent flattening and density modification using the selenomethionine-containing and native data sets. Cycles of SOLVE–RESOLVE routine successfully traced and built 665 residues of the expected 1164, with an R_{factor} 43% and merit figure of 0.55. The remaining model was built manually with COOT followed by refinement with REFMAC [18]. Water molecules were added by COOT and the structure was validated using PROCHECK [30]. All pictures were created using PyMOL [31].

3. Results

3.1. Crystalline packing and overall structure

Both the selenomethionine and the native Xylellain crystals belong to the P1 space group, with similar unit cell parameters, containing four Xylellain monomers (chains A–D) (Fig. 1). A contact surface of 2018 $Å^2$ between chains A/B and C/D is stabilized by polar contacts involving a minimum of 19 residues of each monomer, 13 hydrogen bonds and 6 salt bridges. In this crystal packing the active site and the propeptide domain belong to the dimer interface in the A/B and C/D chains. The atomic contacts between chains A/B with C/D is characterized by fewer interactions.

The adopted residue numbering initiates at Met1 from Xylellain full length sequence (GenBank accession No. AE003869, locus_tag XF_0156, complement of position 159327–160142). Final refinement shows 1435 waters molecules and a ribonucleotide (UDP) molecule at each chain.

The first 22 and 23 N-terminal amino acid residues of chains A–D and B–C, respectively, representing the pET28a-derived sequence that include the hexahistidine-tag and thrombine site,

Table 1
Crystallographic data collection and refinement statistics.

Data set	Native	SelMet
Space group	*P*1	*P*1
Cell dimensions		
a, *b*, and *c* (Å)	55.089, 69.31, 82.36	55.15, 69.21, 82.27
α, β, and γ (Å)	75.86, 75.43, 66.51	76.15, 75.67, 66.64
Wavelength (Å)	1.541	0.972
Resolution (Å)	78.66–1.65 (1.74–1.65)	50.00–1.83 (1.91–1.83)
R_{merge}[a] (%)	5.4 (40.0)	5.4 (40)
I/σ (I)	9.9 (1.9)	15.1 (2.5)
Completeness (%)	91.0 (86.7)	99.5 (97.8)
Redundancy	8.1 (4.4)	4.8 (3.1)
No. of reflections observed	117070	86,291
No. of reflections unique	16,308	7651
Refinement		
Number of atoms	8578	–
R_{work}/R_{free} (%)	16.85/21.59	–
r.m.s.d.		–
Bond angles (°)	2.082	
Bond lengths (Å)	0.025	–
Ramachandran analysis (%/No.)		–
Most favored	88.7/800	
Additional allowed	10.9/98	–
General allowed	0.4/4	–

Numbers in parentheses represent the highest resolution bin.

[a] $R_{merge} = \sum_h \sum_l / I_{hl} - \langle I_h \rangle | \sum_h \sum_l \langle I_h \rangle$, where I_{hl} is the intensity of the *l*th observation of reflection *h* and $\langle I_{hl} \rangle$ are the mean intensity of the *h* reflections. R_{merge} is computed over all *l* observations of *h* reflection.