

## Accepted Manuscript

Highly stable single-strand-specific 3'-nuclease/nucleotidase from *Legionella pneumophila*

Mária Trundová, Tomáš Koval', Raymond J. Owens, Karla Fejfarová, Jarmila Dušková, Petr Kolenko, Jan Dohnálek



PII: S0141-8130(18)30276-9  
DOI: doi:[10.1016/j.ijbiomac.2018.03.113](https://doi.org/10.1016/j.ijbiomac.2018.03.113)  
Reference: BIOMAC 9335

To appear in:

Received date: 19 January 2018  
Revised date: 19 March 2018  
Accepted date: 21 March 2018

Please cite this article as: Mária Trundová, Tomáš Koval', Raymond J. Owens, Karla Fejfarová, Jarmila Dušková, Petr Kolenko, Jan Dohnálek , Highly stable single-strand-specific 3'-nuclease/nucleotidase from *Legionella pneumophila*. The address for the corresponding author was captured as affiliation for all authors. Please check if appropriate. Biomac(2017), doi:[10.1016/j.ijbiomac.2018.03.113](https://doi.org/10.1016/j.ijbiomac.2018.03.113)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Highly stable single-strand-specific 3'-nuclease/nucleotidase from *Legionella pneumophila***

Mária Trundová<sup>a\*</sup>, Tomáš Koval<sup>a\*</sup>, Raymond J. Owens<sup>b</sup>, Karla Fejfarová<sup>a</sup>, Jarmila Dušková<sup>a</sup>, Petr Kolenko<sup>a</sup>, Jan Dohnálek<sup>a</sup>

<sup>a</sup>Laboratory of Structure and Function of Biomolecules, Institute of Biotechnology of the Czech Academy of Sciences, Biocev, Průmyslová 595, 25250 Vestec, Czech Republic

<sup>b</sup>OPPF-UK, The Research Complex at Harwell, Rutherford Appleton Laboratory, Oxfordshire, UK

\*These authors contributed equally.

E-mail addresses of all authors: MT: maria.trundova@ibt.cas.cz, TK: tomas.koval@ibt.cas.cz, RO: ray@strubi.ox.ac.uk, KF: karla.fejfarova@img.cas.cz, JDu: jarmila.duskova@ibt.cas.cz, PK: petr.kolenko@ibt.cas.cz, JDo: dohnalek@ibt.cas.cz.

*Corresponding author*

Jan Dohnálek, Laboratory of Structure and Function of Biomolecules, Institute of Biotechnology of the Czech Academy of Sciences, Biocev, Průmyslová 595, 25250 Vestec, Czech Republic, tel. +420 325 873 758, fax: +420 325 873 710, e-mail: dohnalek@ibt.cas.cz, <http://www.ibt.cas.cz/vyzkum/laboratore/laborator-struktury-a-funkce-biomolekul/>

*Running title*

3'-nuclease from *L. pneumophila*

*Abbreviations*

3'-AMP, adenosine 3'-monophosphate; 5'-AMP, adenosine 5'-monophosphate; CD, circular dichroism; DTT, 1,4-dithiotreitol; EDTA, ethylenediaminetetraacetic acid; IEF, isoelectric focusing; LCV, *Legionella*-containing vacuole; Lp, *Legionella pneumophila*; Lpn1, *Legionella pneumophila* nuclease 1; MALDI, matrix-assisted laser desorption/ionization; MALDI-TOF, matrix-assisted laser desorption/ionization – time-of-flight; MWCO, molecular weight cut off; NBS1, nucleoside binding site 1; Ni-NTA, nickel-nitrilotriacetic acid; Zn-NTA, zinc-nitrilotriacetic acid; OPPE, Oxford protein production facility; PBST, phosphate buffered saline with Tween 20; PDB, Protein Data Bank; R-ABN1, recombinant arabis bifunctional nuclease 1; R-HBN1, recombinant hop bifunctional nuclease 1; SEC, size exclusion chromatography; SDS-PAGE, sodium dodecyl sulfate – polyacrylamide gel electrophoresis; T2SS, type II secretion system; TBN1, tomato bifunctional nuclease 1.

*Enzymes*

Lpn1 – 3'-nuclease/nucleotidase, EC 3.1.30.1

*Conflicts of Interest*

Declarations of interest: none.

Funding: We acknowledge support of this work by the Ministry of Education, Youth and Sports of the Czech Republic (projects LG14009; support of CIISB-Biocev, Biophysics, Advanced Mass Spectrometry, LM2015043), by the Czech Science Foundation (15-05228S), by the ERDF fund (CZ.1.05/2.1.00/19.0390, CZ.02.1.01/0.0/0.0/16\_013/0001776), by Instruct project no. 1058, and by institutional support RVO 86652036.

## Abstract

The Gram-negative bacterium *Legionella pneumophila* is one of the known opportunistic human pathogens with a gene coding for a zinc-dependent S1–P1 type nuclease. Bacterial zinc-dependent 3′-nucleases/nucleotidases are little characterized and not fully understood, including *L. pneumophila* nuclease 1 (Lpn1), in contrast to many eukaryotic representatives with in-depth studies available. To help explain the principle properties and role of these enzymes in intracellular prokaryotic pathogens we have designed and optimized a heterologous expression protocol utilizing *E. coli* together with an efficient purification procedure, and performed detailed characterization of the enzyme. Replacement of Ni<sup>2+</sup> ions by Zn<sup>2+</sup> ions in affinity purification proved to be a crucial step in the production of pure and stable protein. The production protocol provides protein with high yield, purity, stability, and solubility for structure-function studies. We show that highly thermostable Lpn1 is active mainly towards RNA and ssDNA, with pH optima 7.0 and 6.0, respectively, with low activity towards dsDNA; the enzyme features pronounced substrate inhibition. Bioinformatic and experimental analysis, together with computer modeling and electrostatics calculations point to an unusually high positive charge on the enzyme surface under optimal conditions for catalysis. The results help explain the catalytic properties of Lpn1 and its substrate inhibition.

### Keywords

S1–P1 nuclease; *Legionella*; *Escherichia coli* expression

## 1. Introduction

Gram-negative bacterium *Legionella pneumophila* (Lp) causing the Legionnaire's disease [1-2] is capable of infecting human cells, in particular alveolar macrophages [3]. Its natural hosts are mostly amoebae in freshwater environment [4]. Lp subverts the molecular and regulatory pathways of the host cell via more than 300 effectors identified so far [5]. Two life stages of Lp were identified: in the replication stage the bacterium resides and replicates within the modified intracellular compartment – the *Legionella*-containing vacuole (LCV) – of a host cell and utilizes the life cycle mechanisms of the host cell to secure nutrients for its proliferation while suppressing the natural apoptotic and defense mechanisms of the host cell [5]. Upon starvation Lp transforms into an infectious form and develops motility. In this form Lp is capable of survival in fresh water environment for up to several weeks [6].

LCV recruits host cell's endoplasmic reticulum and escapes degradation mechanisms and intracellular immunity pathways. Lp then releases another group of effectors, enzymes and regulatory molecules to secure survival of the host cell and scavenging of nutrients (mainly amino acids). Upon nutrients exhaustion the stringent response leads to transformation of Lp from the replication form into the motile infectious form, shutting down many of its metabolic pathways. Eventually the host cell is burst and the parasite cells are released.

At the time of host cell attack the effectors are translocated into the host cell by the Dot/Icm type IV secretion system [5,7]. 60 or more proteins connected with the replication stage of the bacterium are transported by the Lp type II secretion system (T2SS) [8]. While the former group of effectors play a crucial role during the host cell transformation, the latter are indispensable for pathogen survival and replication in LCV. The majority of nutrients transported into LCV is apparently formed by amino acids but Lp also employs systems for retrieval of nucleic acids, lipids, and other molecules, such as the PhtC and PhtD proteins involved in the thymidine salvage pathway of Lp [9,10].

Extracellular nuclease activity of Lp was reported previously [11]. A mutant strain with affected T2SS had significantly decreased RNase activity but DNase activity remained unaffected. A more recent study into T2SS-transported effectors of Lp identified type 2 ribonuclease [8]. This enzyme is likely responsible for the ribonuclease activity reported earlier by Aragon et al. [11] All the known species of the genus *Legionella* possess a gene for 3'-nuclease/nucleotidase of the S1-P1 family, in *Legionella pneumophila* Philadelphia 1 represented by gene Lpg1570 (UniprotKB, Legiolist – Institut Pasteur).

This protein is missing from mammalian genomes yet present in some bacterial and protozoan pathogens of mammals. Interestingly, also other human pathogens, such as *Leishmania major*, *Leishmania donovani*, *Trypanosoma brucei*, and other protozoan parasites possess or even rely on



such zinc-dependent 3'-nucleases/nucleotidases (for more details see [12]). The gene Lpg1570 for *Legionella pneumophila* nuclease 1 (Lpn1) is transcribed in the replication stage and its transcription is downregulated upon stringent conditions [6].

Lpn1 represents an attractive target from the S1–P1 nuclease family to investigate this enzyme class in bacterial pathogens. Following our previous studies of homologous nucleases from plants and fungi [13–17] we bring the first characterization data of this 3'-nuclease/nucleotidase. S1–P1 nucleases were usually expressed using either the organism of origin or a closely related eukaryotic expression system with the exception of P4 nuclease from *Leishmania infantum* produced in *E. coli* [12]. Our optimized expression and purification protocol enabled production of high quantity and quality of the recombinant enzyme. Characterization of enzymatic and molecular properties uncovered several key features important for the assessment of the role of this enzyme in Lp and for further studies. This work represents the first detailed characterization study of highly stable recombinant zinc dependent bacterial 3'-nuclease/nucleotidase expressed heterologously in *E. coli*.

## 2. Materials and methods

### 2.1 Synthetic Lpn1 gene

The whole gene for Lpn1 (with native signal peptide sequence) from *Legionella pneumophila* subsp. *pneumophila* str. Philadelphia 1, was synthesized by Shanghai Generay Biotech Co. Ltd, according to the GenBank sequence number 52841800.

According to sequence alignment to related S1–P1 nucleases the N-terminus of the mature enzyme was reliably determined as WNAIGH... The N-terminal tryptophan residue is necessary for formation of the zinc-containing active center of the enzyme (Fig. 1). In order to achieve successful expression of recombinant Lpn1, several DNA constructs containing the Lpn1 sequence including or lacking the native signal peptide sequence (855 bp or 774 bp, respectively) were prepared in collaboration with the OPPF center (Oxford Protein Production Facility). The method of “In-Fusion<sup>TM</sup>” cloning was used for the insertion of the Lpn1 gene into a series of the OPPF vectors (Table 1) according to the OPPF manual.

Primer sequences for individual vectors with different tags (N-terminal or C-terminal location, with a preference for C-terminus) are listed in Table 3.

The correct insertion of the Lpn1 gene was verified by sequencing of the vector cloning regions and of the whole inserts as well.

### 2.2 Recombinant Lpn1 expression

Expression of recombinant Lpn1 was tested in several strains of *E. coli*. The expression level was evaluated in *E. coli* BL21 (DE3), *E. coli* BL21(DE3) pLysS, Rosetta<sup>TM</sup> (DE3) pLysS (Novagen), Lemo21 (DE3), SHuffle T7 Express *E. coli* (New England BioLabs), and BL21-CodonPlus(DE3)RIL (Agilent Technologies) in combination with each expression construct.

Transformation of chemically competent *E. coli* cells with prepared expression constructs was performed by the heat shock method [18]. The transformants were selected on LB agar plates with appropriate antibiotics at 37 °C overnight.

Three different bacterial media for recombinant protein production experiments were used: Luria – Bertani Broth (LB broth), Power broth (PB) (Molecular Dimensions), Overnight Express<sup>TM</sup> Instant TB Medium (auto-induction media, Novagen, Merck).

Small scale protein production was used for optimization of expression conditions. In order to gain the best results, the optimization of cultivation temperature, time of production, and density of bacterial culture at the time of the protein induction was performed. The time course of recombinant protein production was monitored using SDS-PAGE.

For expression studies, an overnight culture of *E. coli* expression host cells carrying the expression

construct was diluted at ratio 1:100 into 1 L of fresh LB, PB or auto-induction media supplemented with 50 µg/mL carbenicillin and 30 µg/mL chloramphenicol in the case of the strains: Lemo21 (DE3), *E. coli* BL21(DE3) pLysS, Rosetta<sup>TM</sup> (DE3) pLysS, and BL21-CodonPlus(DE3)RIL. Bacterial cells were grown to OD<sub>600</sub> = 0.6 at 30 °C at a shaking speed of 250 RPM. In the case of LB and PB medium the recombinant protein production was induced by adding isopropyl-β-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Expression level of Lpn1 was monitored at hourly intervals for four hours at 30 °C after induction. The second approach was to decrease the incubation temperature to 20 °C and continue recombinant protein production overnight. Then, bacterial cells were harvested by centrifugation for 20 minutes at 3300 x g, 4 °C, and cell pellets were used immediately or stored at –80 °C until processing. Recombinant protein production was analyzed by SDS-PAGE and Immunoblot.

## **2.3 Purification of recombinant Lpn1 nuclease**

### **2.3.1 Preparation of periplasmic proteins from *E. coli* cells**

After cells harvesting the content of the bacterial periplasmic space was immediately extracted by osmotic shock as follows. Bacterial pellet was carefully resuspended in 50 mM Tris, 20 % sucrose, 1 mM EDTA, pH 7.5 (80 mL buffer solution per gram of wet weight) and incubated on ice for 20 minutes. After centrifugation (3300 x g, 30 minutes at 4 °C), all supernatant was carefully removed and pellet was resuspended in the same volume of ice-cold 5 mM MgSO<sub>4</sub> and kept for 20 minutes on ice bath. The bacterial residues were sedimented by centrifugation (4000 x g, 30 minutes at 4 °C). The supernatant containing periplasmic proteins was filtered through a 0.22 µm filter and further used for Lpn1 nuclease purification.

### **2.3.2 Preparation of soluble fraction of *E. coli* cells lysate**

The bacterial cell pellet was resuspended in binding and washing buffer: 50 mM Tris, 500 mM NaCl, 15 mM imidazole, pH 7.5 at 5 mL per gram of wet weight. Enzymatic lysis of the bacterial cells was performed by adding lysozyme to 0.2 mg/mL, DNase I (Deoxyribonuclease I from bovine pancreas, Sigma-Aldrich) to 20 µg/mL, and 0.5 mL of Protease Inhibitor Cocktail (Sigma-Aldrich) per 10 g of wet weight of bacterial pellet. After incubation on ice bath with stirring for 30 minutes, the bacterial lysate was sonicated on ice using 10 cycles lasting 20 s (at 30 W) with a 40 s cooling period between each burst. Cell debris were removed by centrifugation at 40 000 x g for 30 minutes at 4 °C. Supernatant, containing soluble fraction of bacterial lysate was filtered through a 0.22 µm filter before applying to the affinity chromatography column.

### 2.3.3 Ni<sup>2+</sup> chelate affinity chromatography

Recombinant Lpn1 nuclease fused to His-tag was separated from contaminant bacterial proteins by Ni<sup>2+</sup> chelate affinity chromatography. Soluble cytoplasmic fraction of bacterial lysates or periplasmic fractions of bacterial cells were applied to a HisTrap<sup>TM</sup> FF (1 mL) column (GE Healthcare) and the purification procedure was performed according to the manufacturer's instructions using an ÄKTA purifier (GE Healthcare). Binding and washing buffer composition was 50 mM Tris, 500 mM NaCl, 15 mM imidazole, pH 7.5. Bound proteins were eluted by step gradient of elution buffer: 50 mM Tris, 500 mM NaCl, 500 mM imidazole, pH 7.5. The protein fractions were analyzed by SDS-PAGE and protein concentration was determined spectrophotometrically at 280 nm. Pooled protein fractions were further purified or stored at –80 °C in the presence of 25% (v/v) glycerol.

### 2.3.4 Zn<sup>2+</sup> chelate affinity chromatography

The influence of Ni<sup>2+</sup> ions on quality, stability, and function of Lpn1 nuclease was evaluated by replacing the Ni-NTA purification step by affinity chromatography using a HisTrap column (1 mL) (GE Healthcare), stripped of Ni<sup>2+</sup> ions with 10 column volumes of 20 mM Tris-HCl, 500 mM NaCl, 50 mM EDTA, pH 7.5 and recharged with 0.1 M ZnCl<sub>2</sub> (according to the manufacturer's recommendations). The purification procedure was the same as for the Ni<sup>2+</sup> HisTrap column.

### 2.3.5 Heparin affinity chromatography

Protein fractions from the Ni<sup>2+</sup> or Zn<sup>2+</sup> affinity chromatography were pooled and then loaded on a HiTrap<sup>TM</sup> Heparin HP (1mL, GE Healthcare) column, suitable for separation of DNA-binding proteins. Protein samples were applied to the column after buffer exchange using Microsep<sup>TM</sup> Advance centrifugal devices (10 MWCO) in binding buffer: 50 mM Bis-Tris, pH 6. The column was washed by 10 column volumes of binding buffer and bound proteins were eluted with 10 column volumes of elution buffer containing 50 mM Bis-Tris, 2 M NaCl, pH 6 using a linear gradient. Pooled protein fractions were desalted using Microsep<sup>TM</sup> Advance centrifugal devices (10 MWCO) and their purity was analyzed by SDS-PAGE. Concentration of the purified protein fractions was determined spectrophotometrically (A280).

### 2.3.6 Size exclusion chromatography

Size exclusion chromatography (SEC) was performed using an ÄKTA purifier and a Superdex 75 10/300 GL column (GE Healthcare) calibrated using protein standards. 100 µL of Lpn1 at a concentration of 6 mg/mL was used. SEC was performed in 25 mM Bis-Tris, pH 6.0 with addition

of 150 mM NaCl.

## **2.4 Identification and characterization of recombinant Lpn1**

### **2.4.1 SDS-PAGE and Western blotting**

Aliquots of bacterial cells were mixed with 2x loading buffer (0.126 M Tris-HCl, 20% glycerol, 4% SDS, 0.005% bromophenol blue, 0.1 M DTT, pH 6.8) and heated to 98 °C for 3 minutes. Bacterial lysates were then separated on a NuPage<sup>TM</sup> 4-12% Bis-Tris Gel (Invitrogen, Thermofisher Scientific). The composition of the electrophoresis buffer was 0.05 M MES, 0.05 M Tris Base, 0.1% SDS, 1 mM EDTA, pH 7.3. The gels were run at constant voltage (200 V) at room temperature. The PAGE gels were stained with InstantBlue<sup>TM</sup> (Sigma-Aldrich) and molecular weight of proteins was evaluated using the Color Prestained Protein Standard, Broad Range (New England BioLabs).

In the case of the Western blotting procedure, the separated bacterial proteins were blotted from the SDS-PAGE gels onto a nitrocellulose membrane (Serva) using a Pierce G2 Fast Blotter (Thermofisher Scientific) at 11 V, 1.3 A, at room temperature for 7 minutes with use of the 1-Step Transfer Buffer (Thermofisher Scientific).

### **2.4.2 Immunostaining**

Production of recombinant Lpn1 was evaluated by using antibody against His-tag, located on the C-terminus (or N-terminus) of the protein. The nitrocellulose membrane with blotted bacterial lysates or with periplasmic fractions was blocked with 3% nonfat dry milk in Phosphate buffered saline (Sigma-Aldrich) with 0.1% Tween-20 (PBST) for 1 hour. After washing with PBST, membrane was incubated with monoclonal Anti-HIS Tag antibody (produced in mouse, Sigma-Aldrich) diluted at ratio 1:1000 in MPBST (3% nonfat dry milk in PBST) at room temperature for 1 hour. After thorough rinsing, the membrane was incubated with Anti-Mouse IgG antibody conjugated with horseradish peroxidase (produced in rabbit, Sigma-Aldrich) diluted at ratio 1:500 in MPBST at room temperature for 1 hour. Excess unbound conjugate was washed away with PBST and the reaction was visualized by adding the chromogenic TMB Enhanced One Component HRP Membrane Substrate (3, 3', 5, 5'-tetramethylbenzidine, Sigma-Aldrich).

### **2.4.3 Confirmation of Lpn1 identity by mass spectrometry**

Twenty micrograms of Lpn1 protein were transferred to 20 mM NH<sub>4</sub>HCO<sub>3</sub> buffer, pH 7.8 using Amicon Ultra 0.5 mL centrifugal filters with MWCO 3 kDa (Merck). One µL of Lpn1 protein at concentration 10 pmol/µL was applied on several spots of the 384 stainless steel MALDI target (Bruker Daltonics) and let dry at room temperature. The spots were covered by 2,5-

dihydroxyacetophenone matrix (Bruker Daltonics) for MALDI-TOF analysis of intact protein performed by an Ultraflex III TOF-TOF mass spectrometer (Bruker Daltonics) and by 1,5-diaminonaphthalene matrix (Sigma-Aldrich) for MALDI in-source decay analysis, which was performed by a 15T solariX XR mass spectrometer (Bruker Daltonics). The MALDI in-source decay data were processed by the Data Analysis 4.0 software (Bruker Daltonics) and the analysis was carried out using BioTools 3.2 (Bruker Daltonics). The observed c-type fragments confirmed the primary amino acid sequence of Lpn1.

#### **2.4.4 Determination of experimental pI**

The experimental pI of Lpn1 nuclease was determined by isoelectric focusing (IEF) using an XCell SureLock™ mini-cell electrophoresis system, pre-cast 5% polyacrylamide Vertical Novex® IEF Mini Gels pH 3–10, and IEF Marker 3–10 (Invitrogen, Thermofisher Scientific). Electrophoresis was performed according to the manufacturer's instructions.

#### **2.4.5 Circular dichroism and melting temperature**

Circular dichroism (CD) spectrum was recorded using a Chirascan™-plus spectrometer (Applied Photophysics) and a 0.1 cm path-length quartz cell. Spectrum was recorded over the wavelength range of 195–260 nm in steps of 1 nm at room temperature. Concentration of the sample was 0.2 mg/mL. Sample was diluted from the stock solution (concentration of 2 mg/mL in 25 mM Bis-Tris, pH 6.0 with addition of 50 mM NaCl) by distilled water. The CD signal was expressed as ellipticity and the resulting spectra were buffer subtracted. Secondary structure composition was analyzed using the CDNN program provided with the Chirascan CD spectrometer.

Melting temperature ( $T_m$ ) was determined by the change in the circular dichroism signal at the wavelength of 222 nm. Changes were recorded in the temperature range of 20–94 °C with 12 s measurement at each one-degree step and with a speed of temperature increase of 2 °C per minute.

### **2.5 Activity assays**

#### **2.5.1 Nuclease activity assay**

Nuclease activity of Lpn1 was measured towards calf thymus DNA (Sigma-Aldrich) and RNA from baker's yeast (Sigma-Aldrich). Single-strand nuclease activity was evaluated by using heat-denatured calf thymus DNA (ssDNA) as substrate. Measurements of activity dependence on pH and temperature (in the case of RNA) were performed with a total volume of the reaction mixture of 100 µL containing Lpn1 (1 µg/mL in the case of dsDNA and 0.5 µg/mL in the cases of ssDNA and RNA) and 200 µg/mL dsDNA, ssDNA or RNA. The enzymatic reactions were incubated at 37 °C

for 15 minutes (in the case of ssDNA and RNA substrates) and for 50 minutes in the case of dsDNA. The reaction was stopped by adding 250  $\mu$ L of isopropanol and 10  $\mu$ L of 3 M sodium acetate, pH 5.2. After stirring by vortex the reactions were incubated at  $-20^{\circ}\text{C}$  for 30 minutes. Precipitated undigested substrate was removed by centrifugation (22 000  $\times$  g, 40 minutes,  $4^{\circ}\text{C}$ ) and absorbance of the supernatant was measured at 260 nm. One unit (U) of activity was defined as change of absorbance of 0.001 in 1 cm path per 1 minute per 1  $\mu$ g of enzyme in the total volume of reaction mixture of 100  $\mu$ L [19].

All enzymatic reactions were performed in triplicates including separate background readings for each reaction condition. The influence of pH on the enzymatic activity was analyzed towards each substrate in pH range 3.0–9.0 (for buffer composition see Table 4).

Kinetic parameters were determined by performing reactions containing 0.5  $\mu$ g/mL of Lpn1 in the cases of ssDNA and RNA substrates or 1  $\mu$ g/mL in the case of dsDNA with the following substrate concentration ranges: ssDNA (15.6–1000  $\mu$ g/mL, reaction buffer 50 mM Bis-Tris, 50 mM NaCl, pH 6), RNA (15.6–1000  $\mu$ g/mL, reaction buffer 50 mM Bis-Tris, 50 mM NaCl, pH 6.5), and dsDNA (10–1000  $\mu$ g/mL, reaction buffer 50 mM Bis-Tris, 50 mM NaCl, pH 6). Assay mixtures with ssDNA and RNA were incubated for 15 minutes, whereas with dsDNA for 50 minutes.

Measurements were performed at least for 7 concentrations for each substrate and in triplicates. For all substrates the background reading for each concentration point was evaluated. Kinetic parameters were analyzed using the substrate inhibition equation as implemented in the GraphPad Prism7 software (GraphPad Software, USA).

The temperature optimum of Lpn1 was determined for RNA as substrate (reaction buffer 50 mM Bis-Tris, 50 mM NaCl, pH 6.5) performing the above described assay at 10 points between  $0^{\circ}\text{C}$  and  $80^{\circ}\text{C}$ .

### 2.5.2 Nucleotidase activity

Lpn1 nucleotidase activity was measured towards 3'-AMP and 5'-AMP as substrates. Reactions contained 0.1  $\mu$ g/mL of Lpn1 (0.005  $\mu$ g per 50  $\mu$ L of total volume) and 2 mmol of 3'-AMP or 5'-AMP in a total volume of 50  $\mu$ L. The amount of released phosphate was analyzed according to the method of Lin and Morales [20] by spectrophotometry at  $\lambda = 850$  nm using a microplate reader CLARIOStar (BMG LABTECH).

Determination of the pH optimum was performed in buffers covering pH range 4–9 according to Table 4. The influence of pH in this pH range on the applicability of the above method for determination of pH optimum is not significant, as verified by independent control measurements in triplicates. For determination of kinetic parameters, enzymatic reactions were performed with 0.1  $\mu$ g/mL Lpn1 and with 3'-AMP substrate in concentration range 0.0625 to 8 mM (8 concentration

points). The measurements were carried out in 50 mM Bis-Tris, 50 mM NaCl, pH 7.0 at 37 °C for 5 minutes, in triplicates. The obtained results were expressed as the number of  $\mu\text{mol}$  of released phosphate based on a calibration curve ( $\text{Na}_2\text{HPO}_4$  buffer, pH 7 in range 0.0039–0.5 mM). Kinetic parameters were calculated using the substrate inhibition equation of the GraphPad Prism7 software (GraphPad Software, USA).

### 2.5.3 Effects of $\text{Zn}^{2+}$ and EDTA on ssDNase activity

Lpn1 at concentration 0.5  $\mu\text{g/mL}$  was incubated with 1 mM EDTA or 5 mM  $\text{Zn}^{2+}$  on wet ice for 15 minutes and then the enzymatic assay with ssDNA (200  $\mu\text{g/mL}$ ) as substrate was performed as described above. The measurements were performed in triplicates. Total activities of the untreated and EDTA and  $\text{Zn}^{2+}$  pre-incubated enzyme samples were compared.

### 2.5.4 Phospholipase activity

The phospholipase C-like activity of Lpn1 nuclease was tested by the assay according to Kurioka and Liu [21]. The principle of this method lies in testing the ability of Lpn1 to cleave L- $\alpha$ -phosphatidylcholine (L- $\alpha$ -lecithin) into 1,2-diglyceride and choline phosphate. Reaction mixture of a total volume of 100  $\mu\text{L}$  containing reaction buffer (50 mM Bis-Tris, 50 mM NaCl, pH 7.0), 10  $\mu\text{g}$  of Lpn1, 10  $\mu\text{mol}$  of L- $\alpha$ -phosphatidylcholine, and 0.2 unit of alkaline phosphatase from *E. coli* (Sigma-Aldrich, 1 unit hydrolyzes 1.0  $\mu\text{mol}$  of p-nitrophenyl phosphate per minute at pH 10.4 and 37 °C) was incubated at 37 °C for 2 hours. Inorganic phosphate released by alkaline phosphatase was analyzed according to Lin and Morales [20]. The reaction was stopped by adding 125  $\mu\text{L}$  of reagent A (0.1 M L-ascorbic acid, 0.5 M  $\text{Cl}_3\text{CCOOH}$ ). After thorough mixing, 25  $\mu\text{L}$  of reagent B (10 mM  $(\text{NH}_4)_6\text{Mo}_7\text{O}_{24} \cdot 4\text{H}_2\text{O}$ ) and 62.5  $\mu\text{L}$  of reagent C (0.1 M sodium citrate, 0.2 M  $\text{NaAsO}_2$ , 10% acetic acid) was added. Blue staining formed by the reaction of the compounds with inorganic phosphate was measured spectrophotometrically at  $\lambda = 850 \text{ nm}$ .

### 2.6 Bioinformatic analysis, molecular modeling, and electrostatics calculations

Sequence and signal peptide analysis and alignments were performed with the use of the following software tools and servers: BLAST [22], Clustal Omega [23], GeneDoc [24], PSORTb [25], MultiLoc2 [26], SignalP [27], Phobius Pred [28], TargetP [29], PrediSi (<http://www.predisi.de/>) [30], Pilfind [31], and ESPript 3.0 (<http://esprict.ibcp.fr>) [32]. Signal peptide analysis queries included both prokaryotic and eukaryotic targeting.

Model of Lpn1 was generated using the SwissModel server [33] and the structure of S1 nuclease as a template (PDB id 5FBA, sequence identity 19%) [17]. Partial charges were assigned using the



PropKa software [34] at pH optimal for a given enzyme. Partial charge of  $\text{Zn}^{2+}$  ions in the PropKa output was adjusted to +2. Parameter files were prepared using PDB2PQR [35] on the PDB2PQR server. Distribution of electrostatic potential was determined by solving the Poisson-Boltzmann equation using APBS [36] and displayed with PyMOL (Schrödinger, LLC).

### 3. Results

#### 3.1 Cloning and expression studies

To test the best expression strategy two different approaches were designed. First, the gene coding for Lpn1 (Lpg1570, UniProtKB: Q5ZV70) with the native signal peptide sequence was amplified by PCR and the 855 bp product was cloned into the vector pOPINE [37]. In the second approach the Lpn1 gene fragment of 774 bp, lacking the native signal peptide sequence (81 bp, 27 aa, Fig. 1), was cloned into selected vectors provided by OPPF (Table 1).

While pOPINE and pOPINS are vectors for cytoplasmic recombinant protein expression, the other vectors used in this study were designated for the periplasmic targeting of heterologous proteins.

Successful expression of soluble Lpn1 was detected only in *E. coli* Lemo21 cells transformed with the pOPINMalE–Lpn1 plasmid. Several transformed bacterial colonies were tested for Lpn1 production in a small scale experiment. The bacterial clones expressing the 30 kDa Lpn1 nuclease were propagated on agar plates containing 50 µg/mL carbenicillin, 30 µg/mL chloramphenicol, and 1% glucose. However, passaging of the bacterial culture on agar plates and long-term storage at 4 °C led to loss of Lpn1 expression. It was necessary to optimize several aspects of Lpn1 expression to maximize the yield of the soluble recombinant nuclease. The best results were achieved under the following conditions. Bacterial pre-culture was prepared by inoculation of bacterial cells containing pOPINMalE carrying the Lpn1 gene from glycerol stock into PB medium with 50 µg/mL carbenicillin, 30 µg/mL chloramphenicol, and 1% glucose and cultivated at 30 °C at a shaking speed of 200 RPM for 12 hours. Fresh overnight culture was diluted at ratio 1:100 to the final volume of 1 L of Overnight Express<sup>TM</sup> Instant TB containing 50 µg/mL carbenicillin and 30 µg/mL chloramphenicol and kept at 30 °C at a shaking speed of 250 RPM. When the optical density of bacterial culture reached 0.6, temperature was decreased to 20 °C and the incubation at 250 RPM continued overnight.

### 3.2 Purification

Despite of the periplasmic space targeting with the vector pOPINMalE the presence of recombinant Lpn1 was confirmed neither in the crude extract from the periplasmic space nor in the subsequent purification attempts (examined by SDS-PAGE and immunoblot assay).

Therefore, purification was performed from the whole bacterial cell lysate, in two steps: metal affinity chromatography and heparin-based affinity chromatography. The protein concentration in eluted fractions from the Ni-NTA step was limited due to a decreased solubility of the protein with an upper limit of 1 mg/mL. To overcome this difficulty, the purification procedure was repeated with a HisTrap column charged with  $\text{Zn}^{2+}$  ions instead of  $\text{Ni}^{2+}$ . This approach led to improved solubility of Lpn1. The second step of purification was performed using a heparin column and purified recombinant nuclease was concentrated to 5 mg/mL without precipitation.

The yield of recombinant protein was estimated to be between 5 and 8 mg per 1 L of bacterial culture. Purity and molecular mass were assessed by SDS-PAGE where a band was observed corresponding to a MW of about 30 kDa (theoretical MW 29.8 kDa). The sample quality was sufficient for further biophysical, functional, and structural studies (Fig. 2).

### 3.3 Identity and biophysical properties

Mass spectrometry analysis of proteolytic fragments of Lpn1 after Zn-NTA affinity chromatography confirmed the identity of the protein in its mature form (*i.e.* lacking the N-terminal signal sequence), with N-terminal residues WNAIGH. The experimentally determined isoelectric point of Lpn1 confirmed a high value above 9.5, in agreement with the theoretical value of pI of 9.6 (full mature sequence). Secondary structure of Lpn1 was analyzed using circular dichroism (CD). Lpn1 is mainly  $\alpha$ -helical (Fig. 8A) and contains roughly 70% of  $\alpha$ -helices and 5% of  $\beta$ -sheets. The melting temperature was determined by observing changes in its CD spectrum at 222 nm and it is estimated to be 73 °C (Fig. 8B).

The oligomeric state of Lpn1 was analyzed using size exclusion chromatography (SEC). The elution volume of the peak corresponding to Lpn1 suggests a MW of about 42 kDa and so Lpn1 is either in monomeric or dimeric state in solution (Fig. 9). This uncertainty can also be caused by the possible interaction of Lpn1 with the SEC Superdex medium as Lpn1 carries a high positive charge and possibly hydrophobic residues on its surface.

### 3.4 Activity assays

Substrate specificity and kinetic parameters of Lpn1 were evaluated using substrates RNA, ssDNA, and dsDNA. A comparison of kinetic profiles (Fig. 3) suggests a preference for RNA as substrate compared to ssDNA, while activity on dsDNA reaches only about 8% of the ssDNA activity. Strong substrate/product inhibition was observed for all types of substrate. 3'-nucleotidase activity was confirmed and its kinetic parameters were determined (Table 2 and Fig. 3) with a weaker substrate/product inhibition effect. 5'-nucleotidase activity and phospholipase activity was not observed.

### 3.5 pH and temperature optima and metal dependence

In the case of RNA as substrate the enzyme activity reached a maximum around pH 7.0 (Fig. 4A). The pH optimum for ssDNA was around 6.0 (Fig. 4A). 3'-AMP was cleaved by Lpn1 in a broad range of pH with a maximum activity in pH range 6.0–7.0 (Fig. 4B). Lpn1 activity towards RNA in presence of 1 mM EDTA decreased to less than 50%. Addition of 5 mM ZnCl<sub>2</sub> to the purified enzyme without any EDTA treatment had no effect on activity (Suppl. Fig. S1). The RNase activity of Lpn1 has a temperature optimum around 65 °C (Fig. 4C).

### 3.6 Bioinformatic analysis

An exhaustive analysis of the Lpn1 sequence, and especially of its N-terminal signal peptide, using a number of available web-based services, led to the conclusion that the native signal peptide targets Lpn1 almost certainly for secretion. No particular target compartment (extracellular, periplasmic, host cell, LCV) can be identified based on the particular features of the signal sequence.

## 4. Discussion

### 4.1 Periplasmic targeting overcomes Lpn1 toxicity at expression and secures correct processing of N-terminus

Bacterial expression of toxic proteins represents a challenging obstacle in production of high amounts of recombinant proteins for basic research, such as structural studies and biophysical characterization, but also for applications. Periplasmic targeting promises to be an effective way of solving both the toxicity problem and the need to preserve the precise processing of the N-terminus of a recombinant protein [38, 39].

Even if several constructs with cytoplasmic and periplasmic targeting and some with fusion proteins were prepared to obtain soluble Lpn1 nuclease, only periplasmic targeting proved successful. Both major *E. coli* periplasmic transport routes for recombinant proteins SecB [40] and SRP [41] were tested with the use of the relevant signal peptides (Table 1). Successful expression of a reasonable amount of Lpn1 was detected only in the case of the construct pOPINMale–Lpn1, carrying the periplasmic targeting signal peptide male originating from maltose binding protein (MBP). MBP protein translocation occurs via the SecB pathway posttranslationally [42, 43]. It was necessary to optimize the combination of vector type and production strain. The standard purification protocol for products in periplasm did not provide the target protein. Only the entire cell harvesting led to sufficient amounts of protein. The designed construct of the signal sequence yielded successfully processed mature enzyme with the key Trp residue, highly conserved in this enzyme family, at the N-terminus. The involvement of the SecB pathway and the presence of the processed N-terminus implicate localization of the product within the periplasm of *E. coli*. The failure of the standard periplasmic purification protocol (and of several tested modifications) to isolate and detect Lpn1, however, indicates that it was not available in the periplasmic space as a solubilized product. Interestingly, Rahbarnia et al. [44] reported successful production of *L. infantum* nuclease P4 in *E. coli* using periplasmic targeting, however, with utilization of the DsbA signal sequence. In our expression tests the exploitation of the male signal sequence yielded significantly better results than the DsbA/SRP pathway.

The strong positive surface electrostatic potential of Lpn1 (see further details in Discussion) suggests that, similarly as in other ribonucleases [45], high affinity and binding of Lpn1 to membrane surface may occur. This offers a plausible explanation of the periplasmic targeting of the product, failure of the periplasmic purification protocol and success of the more vigorous protocol for whole cells.

A key aspect of the optimized purification protocol was the replacement of Ni-NTA by Zn-NTA, which significantly improved solubility of the purified protein. This could be explained by two independent effects – one being the use of metal ions native to this protein and another being a

decreased aggregation via the C-terminal His-tag with  $\text{Zn}^{2+}$  ions compared to  $\text{Ni}^{2+}$ . As a result, we present a method providing an easy route for bacterial overexpression of this type of nucleases, otherwise toxic for cells.

#### 4.2 Lpn1 is a single-strand-specific endonuclease with significant substrate inhibition

The measured kinetic profiles of Lpn1 towards substrates of various types show that Lpn1 is a single-strand-specific nuclease preferring RNA to ssDNA as substrate with the optimal pH for activity being around 7.0 for RNA and 6.0 for ssDNA (Fig. 4A). Lpn1 also cleaves dsDNA, however at approximately 13x lower rate than ssDNA. Such activity profile is common also for other single-strand-specific S1–P1 nucleases. As a consequence of the pH optimum difference Lpn1 is much more active towards RNA at neutral pH and towards ssDNA at pH 6.0. Thus depending on the conditions the enzyme can be considered an RNase or a DNase. This behavior is similar to that of M1 nuclease from *Mesorhizobium loti* [46], however, the points of M1 pH optima are shifted towards the acidic region: 6.0 for RNA and 5.0 for ssDNA. P1 nuclease from *Penicillium citrinum* prefers RNA as substrate with a pH optimum of 5.4 [47]. S1 and mung bean nuclease are considered acidic single-strand-specific DNases [48]. There are also examples of fungal and plant S1–P1 nucleases with pH optima of RNase and/or DNase activity around neutral or higher pH, such as *Lentinus edodes* Le3 [49], barley nuclease BEN1 [50], or celery nuclease CEL I [51]. Lpn1 also performs as a 3'-nucleotidase with weak substrate/product inhibition.

A majority of characterized S1–P1 nucleases belong to eukaryotic organisms (TBN1, P1, S1, and other) and rely on N-glycosylation as a stabilization strategy (AtBFN2 [52], TBN1 [15]). Here we demonstrated that nuclease Lpn1 of bacterial origin expressed in *E. coli*, and naturally lacking glycosylation, maintains its high stability and function.

A striking feature is the strong inhibition of RNase and DNase activity at higher substrate/product concentrations (Fig. 3). A comparison of Lpn1 enzyme kinetics with data for other nucleases of the same family is not possible as none of the previous studies, except for our work on S1 nuclease [17], reported full kinetic profiles. A weak effect of substrate inhibition was observed in the case of S1 nuclease for mutants and ssDNA as substrate, whereas the wild type of Lpn1 suffers significant substrate/product inhibition with all the studied substrates (ssDNA, RNA, dsDNA, and 3'-AMP).

### 4.3 Extreme electrostatics of Lpn1 helps explain its kinetics

The value of theoretical pI of Lpn1 of 9.6 and the experimental value higher than 9.5 distinguish Lpn1 from the other so far characterized S1–P1 type nucleases, mostly with pI around acidic or slightly acidic values (S1 4.2, P1 4.7, M1 5.0, TBN1 5.5, AtBFN2 5.6). The pH optima of Lpn1 (pH 6.0 and 7.0 for DNA and RNA cleavage, respectively) mean that under favorable conditions for catalysis the enzyme molecule has a significant positive surface charge. The estimated total protein charge is +20 at pH 6.0 and +16 at pH 7.0 (PropKa) [34], calculated on a model of 3D structure of Lpn1 generated on the template of S1 nuclease (PDB id 5FBB), opposed to approximately 0 for S1 nuclease at its pH optimum of 4.0 for ssDNase activity. The overall electrostatic potential distribution calculated for this model (Fig. 5) confirms that under the optimal pH conditions the active site groove of the enzyme is governed by strong positive electrostatic potential. Significant positive electrostatic potential near the active site was also observed in extensive structure-function studies of the plant nuclease TBN1 from the S1–P1 family [13–16], where it was expected to participate mainly in binding of structured nucleic acids.

The exact role of the positive surface charge of Lpn1 cannot be currently explained in the context of the Lp environment or macrophage localization but it helps explain the enzyme kinetics with substrate inhibition (Fig. 3). Within the range of the measured substrates in this study, a stronger affinity to substrate (lower  $K_M$ ) is accompanied by a more significant substrate inhibition; affinity grows from small substrate 3'-AMP, over single-strand substrates towards dsDNA. This behavior can be explained by the surface electrostatic properties of Lpn1. The high positive potential enables good affinity of substrates growing towards structured dsDNA with the exposed negatively charged sugar-phosphate backbone. At the same time it likely hampers product release or may enable non-productive substrate binding. Therefore we propose that the extremely positive electrostatics of Lpn1 is related to the substrate/product inhibition effects in the kinetic profiles.

Ribonucleases belong to a class of molecules utilizing their cationic character to enable cell membrane penetration. Their high positive electrostatic potential enables membrane surface binding and helps their internalization without involvement of any specific receptors [53]. It was also shown that protein modification increasing ribonuclease surface charge led to higher cell wall penetration and cytotoxicity [45]. Taking into account also the above described results of the Lpn1 periplasmic expression and purification, it is likely that Lpn1 naturally binds to membrane surface.

#### 4.4 Lpn1 specificity

Lpn1 possesses 3'-nucleotidase activity and is not able to cleave phosphate from 5'-AMP; this behavior is similar as in many other representatives of the S1–P1 nuclease family, such as TBN1, R-HBN1, R-ABN1, S1, P1, and other. Sequence similarity of the key sites of Lpn1 to other family members, *i.e.* in the nucleobase binding site 1 (NBS1), the zinc cluster, and the positive residue facing the catalytic site (Lys or Arg), is very high. Therefore, we propose that a very similar binding mode of a mononucleotide substrate as in our detailed study of S1 nuclease [17] can be expected. The situation is, however, different in the case of a single-strand oligo- or polynucleotide binding to the active site. The part of the enzyme sequence, which appears to play an important role in all the fully characterized S1–P1 nucleases in the substrate-enzyme interaction behind the catalytic cluster, *i.e.* for binding of nucleotides in positions P1', P2', *etc.*, is “behind” the zinc cluster (in the 5'–3' direction) and juxtaposed to NBS1 with respect to the cluster. Based on the previous structure-function studies it can be assigned to a specific part of the enzyme sequence, in the case of Lpn1 residues IGNN ... KFFQ, numbers 177 to 200 (see Fig. 6), corresponding to a particular position on the enzyme surface. In the previous studies this surface patch of the homologous enzymes was shown to contain the Tyr site in the case of P1 nuclease, the Half-Tyr site in S1 nuclease or the significant positive surface charge patch in the plant nuclease TBN1. The typical substrate binding sites cannot be identified in Lpn1, based on the sequence analysis alone (Fig. 6). This is also shown in Figure 7 by color-coded sequence similarity displayed on the surface of one of the most similar structurally characterized members of the S1–P1 family. While the catalytic center and NBS1 are highly conserved (red and green), the specificity-defining region is not (green and blue). As demonstrated on a small set of amino acid sequences (Fig. 1) there is a significant similarity in the specificity region within the compared legionellae, while a distinct pattern is present near the end of this region in *Stenotrophomonas* and *Pseudomonas* (region MLNTRKL, numbers 162–168 in *P. aeruginosa*). Therefore, within the genus *Legionella* it can be expected that the Lpn1 homologs have similar enzymatic characteristics, with yet unknown underlying mechanism of substrate binding in this region, while a different binding pattern is present in other pathogenic bacteria. It should be noted that the deletion in Lpn1 compared to TBN1 (residues RDCHDQHD, 100–107 in TBN1 numbering), earlier suggested to be important for dsDNA binding (Fig. 7 – below NBS1), is consistent with the prevailing activity of Lpn1 towards single-stranded substrates. This amino acid sequence is partially found in other described nucleases (Fig. 6) but entirely missing in Lpn1. Existence of any specific sites for single-stranded nucleic acid binding remains to be explained by structural studies.



#### 4.5 Lpn1 is a secreted replication stage-related protein

Based on the sequence analysis of available bacterial genomes it can be concluded that a highly homologous enzyme is coded by all known legionellae. The high level of sequence identity (92–100% within 52 genomes, data not shown) suggests that the molecular properties and role of the enzyme are the same in all strains of *Legionella pneumophila*.

From the currently available data the exact localization and role of Lpn1 cannot be determined, however, there are several important factors leading to preliminary conclusions. Given the toxicity of Lpn1 to *E. coli* cells, the neutral pH preference, and the nature of the native signal sequence of Lpn1, it can be proposed that Lpn1 is not active within the replicating Lp cells. Under normal conditions cytoplasmic pH of *E. coli* is maintained between 7.2 and 7.8 [54]. The range of pH at which Lpn1 keeps significant activity is wider for RNA as substrate and covers a rather broad range of about 5.0–8.5, compared to a narrower one for DNA, about 5.0–7.0. If properly processed in *E. coli* cytoplasm Lpn1 should be fully active as an RNase and therefore become toxic by RNA degradation. This explains the intracellular toxicity of Lpn1 for *E. coli* host cells observed in expression experiments. The nature of the signal sequence of Lpn1 suggests its export certainly out of the Lp cell, possibly into the host cell, a human alveolar macrophage in the case of legionellosis, or into LCV. Presence of mature Lpn1 inside the Lp bacterial cell would be toxic for the pathogen for similar reasons as in *E. coli*. Regular pH of human lymphocyte cytoplasm is also maintained around pH 7.0 [55], *i.e.* the internal conditions in macrophage cytoplasm again correspond to the optimal point for Lpn1 RNase activity. The transcriptomic study by Li et al. [6] registered a marked decrease of transcription of many genes of *Legionella pneumophila* upon introduction of stringent conditions after the replication stage. Transcription of the Lpn1 gene Lpg1570 also showed a significant decrease. Therefore it is very likely that Lpn1 is a replication stage-related protein, and given its intracellular toxicity, its expression and/or activity in the host-pathogen interaction must be regulated. Given its strongly cationic character it is possible that this regulation is achieved by membrane surface binding. In such case the enzyme would not be freely available *e.g.* in macrophage cytosol.

The gene Lpg1570 for Lpn1 is found in operon 30005 of *Legionella pneumophila* Philadelphia 1 (Database of prokaryotic operons, NC\_002942:1736886..1742578), which contains seven genes in a reverse strand. None of these is an identified effector in the infection or replication stage [7, 8]. Thus the localization of the gene in this operon is consistent with a role in the standard replication cycle of Lp rather than being directly connected with a modification of the host cell in either stage. There is a thymidine salvage pathway involved in Thy transport and utilization in Lp in the replication stage inside LCV [9]. It can be speculated that Lpn1 may function as an enzyme

securing DNA degradation to provide products for this pathway. The exact role of Lpn1 is, however, still not known and requires further study.

#### 4.6 Lpn1 homologs in other pathogens

While S1–P1 nucleases in plants act in apoptotic processes and fungi such as *Penicillium* or *Aspergillus* are expected to utilize them in nucleotide/nucleoside scavenging in their environment, the role of this enzyme in some protozoa such as *Leishmania* was in part uncovered [56], whereas in parasitizing bacteria it remains unclear. In leishmaniae the S1–P1 type enzyme plays an important role in purine scavenging under purine stress conditions or facilitates interaction with the host [57]. Interestingly, many pathogenic bacteria invading host cells harbor genes for this enzyme and the high degree of sequence similarity suggests a similar function *e.g.* in *Legionella pneumophila*, *Legionella longbeachae*, *Stenotrophomonas maltophilia* or *Pseudomonas aeruginosa*. As follows from the sequence alignment of selected bacterial 3'-nucleases/nucleotidases from known pathogens, these show a high degree of similarity even within bacteria and trypanosomatids (Fig. 1). The main enzyme signatures covering the N-terminal Trp residue, the zinc-coordinating residues, and NBS1 are conserved. The sequences can be grouped by the similarity of the signal peptide to a group of legionellae, a group of *P. aeruginosa* and *S. maltophilia*, and trypanosomatids. In bacteria, this may reflect the different nature of infections caused by legionellae and by *Stenotrophomonas* and *Pseudomonas*, in connection with the nuclease localization. While the former ones infect mammalian cells by invasion and subversion of the molecular mechanisms for successful intracellular parasitic replication, the latter pathogens mainly colonize human organs such as lungs (*Stenotrophomonas*) [58] even if also intracellular invasion was detected (*Pseudomonas*) [59].

#### 5. Conclusion

*Legionella pneumophila* nuclease Lpn1 is the first S1–P1 type nuclease from a human bacterial pathogen successfully expressed in *E. coli*. It is a single-strand-specific nuclease preferring RNA as substrate with activity optimized for neutral or slightly acidic conditions, with a high degree of substrate inhibition. Its unusually high isoelectric point within this enzyme family helps explain the inhibition effects and suggests its binding to cell membrane under standard conditions. The predicted region defining the substrate specificity, derived from other family members, has a new composition which cannot be matched to the known substrate binding patterns or sites.

**Acknowledgements**

The authors wish to thank Drs. Tatsiana Charnavets and Petr Pompach from the Centre of Molecular Structure, BIOCEV for help and support in biophysical techniques and mass spectrometry analysis.

**Author Contributions**

JDo, TK, RJO, and MT planned experiments; TK, MT, and PK performed experiments; TK, JDo, MT, KF, and JDu analyzed data; JDo, MT, and TK wrote the paper; JDo, JDu, and PK other.

**Supplementary material**

Supplementary material is available for the influence of EDTA and zinc on Lpn1 activity.

## References

- [1] D.W. Fraser, T.R. Tsai, W. Orenstein, W.E. Parkin, H.J. Beecham, R.G. Sharrar, J. Harris, G.F. Mallison, S.M. Martin, J.E. McDade, C.C. Shepard, P.S. Brachman, Legionnaires' disease: description of an epidemic of pneumonia, *N. Engl. J. Med.* 297 (1977) 1189-1197.
- [2] J.E. McDade, C.C. Shepard, D.W. Fraser, T.R. Tsai, M.A. Redus, W.R. Dowdle, Legionnaires' disease: isolation of a bacterium and demonstration of its role in other respiratory disease, *N. Engl. J. Med.* 297 (1977) 1197-203.
- [3] M.A. Horwitz, Interactions between macrophages and *Legionella pneumophila*, *Curr. Top. Microbiol. Immunol.* 181 (1992) 265-82.
- [4] J.M. Berjeaud, S. Chevalier, M. Schlusshuber, E. Portier, C. Loiseau, W. Aucher, O. Lesouhaitier, J. Verdon, *Legionella pneumophila*: The Paradox of a Highly Sensitive Opportunistic Waterborne Pathogen Able to Persist in the Environment, *Front. Microbiol.* 7 (2016), <https://doi.org/10.3389/fmicb.2016.00486>.
- [5] E.C. So, C. Mattheis, E.W. Tate, G. Frankel, G.N. Schroeder, Creating a customized intracellular niche: subversion of host cell signaling by *Legionella* type IV secretion system effectors, *Can. J. Microbiol.* 61 (2015) 617-35.
- [6] L. Li, N. Mendis, H. Trigui, S.P. Faucher, Transcriptomic changes of *Legionella pneumophila* in water, *BMC Genomics* 16 (2015), <https://doi.org/10.1186/s12864-015-1869-6>.
- [7] W. Zhu, S. Banga, Y. Tan, C. Zheng, R. Stephenson, J. Gately, Z.Q. Luo, Comprehensive identification of protein substrates of the Dot/Icm type IV transporter of *Legionella pneumophila*, *PLoS One* 6 (2011), <https://doi.org/10.1371/journal.pone.0017638>.
- [8] L. Gomez-Valero, C. Rusniok, C. Cazalet, C. Buchrieser, Comparative and functional genomics of legionella identified eukaryotic like proteins as key players in host-pathogen interactions, *Front. Microbiol.* 2 (2011), <https://doi.org/10.3389/fmicb.2011.00208>.

- [9] M.V. Fonseca, J.D. Sauer, S. Crepin, B. Byrne, M.S. Swanson, The phtC-phtD locus equips *Legionella pneumophila* for thymidine salvage and replication in macrophages. *Infect. Immun.* 82 (2014) 720-730.
- [10] M.V. Fonseca, M.S. Swanson, Nutrient salvaging and metabolism by the intracellular pathogen *Legionella pneumophila*, *Front. Cell. Infect. Microbiol.* 4 (2014), <https://doi.org/10.3389/fcimb.2014.00012>.
- [11] V. Aragon, S. Kurtz, A. Flieger, B. Neumeister, N.P. Cianciotto, Secreted enzymatic activities of wild-type and pilD-deficient *Legionella pneumophila*, *Infect. Immun.* 68 (2000) 1855-1863.
- [12] T. Koval', J. Dohnálek, Characteristics and application of S1–P1 nucleases in biotechnology and medicine, *Biotechnol. Adv.* in press (2017), <https://doi.org/10.1016/j.biotechadv.2017.12.007>.
- [13] J. Dohnálek, T. Koval', P. Lipovová, T. Podzimek, J. Matoušek, Structure analysis of group I plant nucleases, *J. Synchrotron Radiat.* 18 (2011) 29-30.
- [14] T. Koval', P. Lipovová, T. Podzimek, J. Matoušek, J. Dušková, T. Skálová, A. Štěpánková, J. Hašek, J. Dohnálek, Crystallization of recombinant bifunctional nuclease TBN1 from tomato, *Acta Crystallogr. F*67 (2011) 124-128.
- [15] T. Koval', P. Lipovová, T. Podzimek, J. Matoušek, J. Dušková, T. Skálová, A. Štěpánková, J. Hašek, J. Dohnálek, Plant multifunctional nuclease TBN1 with unexpected phospholipase activity: structural study and reaction–mechanism analysis, *Acta Crystallogr. D*69 (2013) 213–226.
- [16] J. Stránský, T. Koval', T. Podzimek, A. Týcová, P. Lipovová, J. Matoušek, P. Kolenko, K. Fejfarová, J. Dušková, T. Skálová, J. Hašek, J. Dohnálek, Phosphate binding in the active centre of tomato multifunctional nuclease TBN1 and analysis of superhelix formation by the enzyme, *Acta Crystallogr. F*71 (2015) 1408-1415.

- [17] T. Koval', L.H. Østergaard, J. Lehmbeck, A. Nørgaard, P. Lipovová, J. Dušková, T. Skálová, M. Trundová, P. Kolenko, K. Fejfarová, J. Stránský, L. Švecová, J. Hašek, J. Dohnálek, Structural and catalytic properties of S1 nuclease from *Aspergillus oryzae* responsible for substrate recognition, cleavage, non-specificity, and inhibition, PLoS ONE 11 (2016), <https://doi.org/10.1371/journal.pone.0168832>.
- [18] A. Froger, J.E. Hall, Transformation of plasmid DNA into E. coli using the heat shock method, J. Vis. Exp. 6 (2007), <https://doi.org/10.3791/253>.
- [19] T. Podzimek, J. Matousek, P. Lipovova, P. Pouckova, V. Spiwok, J. Santrucek, Biochemical properties of three plant nucleases with anticancer potential, Plant Sci. 180 (2011) 343-351.
- [20] T.I. Lin, M.F. Morales, Application of a one-step procedure for measuring inorganic phosphate in the presence of proteins: the actomyosin ATPase system, Anal. Biochem. 77 (1977) 10-17.
- [21] S. Kurioka, P.V. Liu, Improved assay method for Phospholipase C, Appl. Microbiol. 15 (1967) 551-555.
- [22] G.M. Boratyn, C. Camacho, P.S. Cooper, G. Coulouris, A. Fong, N. Ma, T.L. Madden, W.T. Matten, S.D. McGinnis, Y. Merezuk, Y. Raytselis, E.W. Sayers, T. Tao, J. Ye, I. Zaretskaya, BLAST: a more efficient report with usability improvements, Nucleic Acids Res. 41 (2013) W29-33.
- [23] F. Sievers, A. Wilm, D. Dineen, T.J. Gibson, K. Karplus, W. Li, R. Lopez, H. McWilliam, M. Remmert, J. Söding, J.D. Thompson, D.G. Higgins, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, Mol. Syst. Biol. 7 (2011) 539.
- [24] K.B. Nicholas, H.B. Nicholas Jr., D.W. Deerfield II, GeneDoc: Analysis and Visualization of Genetic Variation, Embnew News 4 (1997) 14.
- [25] N.Y. Yu, J.R. Wagner, M.R. Laird, G. Melli, S. Rey, R. Lo, P. Dao, S.C. Sahinalp, M.

- Ester, L.J. Foster, F.S.L. Brinkman, PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes, *Bioinformatics* 26 (2010) 1608-1615.
- [26] T. Blum, S. Briesemeister, O. Kohlbacher, MultiLoc2: integrating phylogeny and Gene Ontology terms improves subcellular protein localization prediction, *BMC Bioinformatics* 10 (2009), <https://doi.org/10.1186/1471-2105-10-274>.
- [27] H. Nielsen, Predicting Secretory Proteins with SignalP, *Methods Mol. Biol.* 1611 (2017) 59-73.
- [28] L. Käll, A. Krogh, E.L.L. Sonnhammer, Advantages of combined transmembrane topology and signal peptide prediction – the Phobius web server, *Nucleic Acids Res.* 35 (2007) W429-432.
- [29] O. Emanuelsson, H. Nielsen, S. Brunak, G. Heijne, Predicting subcellular localization of proteins based on their N-terminal amino acid sequence, *J. Mol. Biol.* 300 (2000) 1005-1016.
- [30] K. Hiller, PrediSi: PREDIction of SIgnal peptides, (2003) <http://www.predisi.de>.
- [31] S. Imam, Z. Chen, D.S. Roos, M. Pohlschröder, Identification of surprisingly diverse type IV pili, across a broad range of gram-positive bacteria, *PLoS One* 6 (2011), <https://doi.org/10.1371/journal.pone.0028919>.
- [32] X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server, *Nucleic Acids Res.* 42 (2014) W320-324.
- [33] M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T.G. Cassarino, M. Bertoni, L. Bordoli, T. Schwede, SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, *Nucleic Acids Res.* 42 (2014) W252-258.

- [34] C.R. Sondergaard, M.H.M. Olsson, M. Rostkowski, J.H. Jensen, Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pKa values, *J. Chem. Theory Comput.* 7 (2011) 2284-2295.
- [35] T.J. Dolinsky, J.E. Nielsen, J.A. McCammon, N.A. Baker, PDB2PQR: an automated pipeline for the setup, execution, and analysis of Poisson-Boltzmann electrostatics calculations, *Nucleic Acids Res.* 32 (2004) W665-667.
- [36] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Electrostatics of nanosystems: application to microtubules and the ribosome, *Proc. Natl. Acad. Sci. USA* 98 (2001) 10037-10041.
- [37] N.S. Berrow, D. Alderton, S. Sainsbury, J. Nettleship, R. Assenberg, N. Rahman, D.I. Stuart, R.J. Owens, *Nucleic Acids Res.* 35 (2007), <https://doi.org/10.1093/nar/gkm047>.
- [38] J. Bogomolova, B. Simon, M. Sattler, G. Stier, Screening of fusion partners for high yield expression and purification of bioactive viscotoxins, *Protein Expr. Purif.* 64 (2009) 16-23.
- [39] G.L. Rosano, E.A. Ceccarelli, Recombinant protein expression in *Escherichia coli*: advances and challenges, *Front. Microbiol.* 5 (2014), <https://doi.org/10.3389/fmicb.2014.00172>.
- [40] H.C. Lee, H.D. Bernstein, The targeting pathway of *Escherichia coli* presecretory and integral membrane proteins is specified by the hydrophobicity of the targeting signal, *P. Natl. Acad. Sci. USA* 98 (2001) 3471-3476.
- [41] C.F. Schierle, M. Berkmen, D. Huber, C. Kumamoto, D. Boyd, J. Beckwith, The DsbA signal sequence directs efficient, cotranslational export of passenger proteins to the *Escherichia coli* periplasm via the signal recognition particle pathway, *J. Bacteriol.* 185 (2003) 5706-5713.
- [42] C.A. Kumamoto, J. Beckwith, Mutations in a new gene, secB, cause defective protein localization in *Escherichia coli*, *J. Bacteriol.* 154 (1983) 253-260.



- [43] P.J. Bassford Jr., Export of the periplasmic maltose-binding protein of *Escherichia coli*, J. Bioenerg. Biomembr. 22 (1990) 401-439.
- [44] L. Rahbarnia, S. Farajnia, B. Naghili, Application of DsbA signal peptide for soluble expression of *Leishmania infantum* P4 nuclease in *E. coli*. Asian J. Anim. Vet. Adv. 7 (2012) 326–333.
- [45] N.K. Sundlass, C.H. Eller, Q. Cui, R.T. Raines, Contribution of electrostatics to the binding of pancreatic-type ribonucleases to membranes, Biochem. 52 (2013) 6304-6312.
- [46] M. Pimkin, C.G. Miller, L. Blakesley, C.A. Oleykowski, N.S. Kodali, A.T. Yeung, Characterization of a periplasmic S1-like nuclease coded by the *Mesorhizobium loti* symbiosis island, Biochem. Biophys. Res. Commun. 343 (2006) 77-84.
- [47] Y. Guo-Qing, L.E. Shi, Y. Yu, T. Zhen-Xing, C. Jian-Shu, Production, purification and characterization of nuclease P1 from *Penicillium citrinum*, Process Biochem. 41 (2006) 1276-1281.
- [48] N.A. Desai, V. Shankar, Single-strand-specific nucleases, FEMS Microbiol. 26 (2003) 457-491.
- [49] H. Kobayashi, N. Inokuchi, T. Koyam, M. Tomita, M. Irie, Purification and characterization of the 2nd 5'-nucleotide-forming nuclease from *Lentinus edodes*, Biosci. Biotechnol. Biochem. 59 (1995) 1169-1171.
- [50] P.H. Brown, T.H. Ho, Biochemical properties and hormonal regulation of barley nuclease, Eur. J. Biochem. 168 (1987) 357-364.
- [51] B. Yang, X. Wen, N.S. Kodali, C.A. Oleykowski, C.G. Miller, J. Kulinski, D. Besack, J.A. Yeung, D. Kowalski, A.T. Yeung, Purification, cloning, and characterization of the CEL I nuclease, Biochem. 39 (2000) 3533-3541.
- [52] C.Y. Ko, Y.L. Lai, W.Y. Liu, C.H. Lin, Y.T. Chen, L.F. Chen, T.Y. Lin, J.F. Shaw, Arabidopsis ENDO2: its catalytic role and requirement of N-glycosylation for function, J. Agric. Food Chem. 60 (2012) 5169-5179.

- [53] R.J. Johnson, T.Y. Chao, L.D. Lavis, R.T. Raines, Cytotoxic ribonucleases: the dichotomy of coulombic forces, *Biochem.* 46 (2007) 10308-10316.
- [54] J.C. Wilks, J.L. Slonczewski, pH of the cytoplasm and periplasm of *Escherichia coli*: rapid measurement by green fluorescent protein fluorimetry, *J. Bacteriol.* 189 (2007) 5601-5607.
- [55] C. Deutsch, J.S. Taylor, D.F. Wilson, Regulation of intracellular pH by human peripheral blood lymphocytes as measured by  $^{19}\text{F}$  NMR, *Proc. Natl. Acad. Sci. USA* 79 (1982) 7944-7948.
- [56] M.B. Joshi, D.M. Dwyer, Molecular and functional analyses of a novel class I secretory nuclease from the human pathogen, *Leishmania donovani*, *J. Biol. Chem.* 282 (2007) 10079–10095.
- [57] A.B. Guimaraes-Costa, T.S. DeSouza-Vieira, R. Paletta-Silva, A.L. Freitas-Mesquita, J.R. Meyer-Fernandes, E.M. Saraiva, 3'-nucleotidase/nuclease activity allows *Leishmania* parasites to escape killing by neutrophil extracellular traps, *Infect. Immun.* 82 (2014) 1732–1740.
- [58] J.S. Brooke, *Stenotrophomonas maltophilia*: an emerging global opportunistic pathogen, *Clin. Microbiol. Rev.* 25 (2012) 2-41.
- [59] R.R. Lovewell, Y.R. Patankar, B. Berwin, Mechanisms of phagocytosis and host clearance of *Pseudomonas aeruginosa*, *Am. J. Physiol. Lung Cell. Mol. Physiol.* 306 (2014) L591-603.

## TABLES

**Table 1**

Expression constructs used for Lpn1 expression. Signal peptides for periplasmic targeting using the SecB and SRP *E. coli* transport pathways are shown. Position of tag is marked with protein terminus C-... or N-...

Expression Vectors	Signal peptide/pathway	Recombinant protein targeting	Tag
pOPINE	Native Lpn1 SP	Cytoplasmic	C-KHIS6
pOPINE	none	Cytoplasmic	
pOPINMaIE	MaIE/SecB	Periplasmic	
pOPINDsba	Dsba/SRP		
pOPINO	OmpA/SecB		
pOPINP	PeIB/SecB		
pOPINTolB	TolB/SRP		
pOPINS	None	Cytoplasmic	N-HIS6-SUMO

**Table 2**

Nuclease and nucleotidase activity of Lpn1.  $v_{\max}$  is given as  $\Delta A_{260}$  of 0.001 per 1  $\mu\text{g}$  of Lpn1 per minute for nuclease activity and as  $\mu\text{mol}$  of released phosphate per 1  $\mu\text{g}$  of Lpn1 per minute for nucleotidase activity.

Substrate	Nuclease activity		
	$K_m$ (mg/mL)	$v_{\max}$	$K_i$ (mg/mL)
RNA	$0.15 \pm 0.04$	$3800 \pm 900$	$0.19 \pm 0.06$
ssDNA	$0.12 \pm 0.02$	$2900 \pm 500$	$0.09 \pm 0.02$
dsDNA	$< 0.01^a$	$230 \pm 70$	$0.07 \pm 0.04$
	Nucleotidase activity		
	$K_m$ (mmol/L)	$v_{\max}$	$K_i$
3'-AMP	$0.7 \pm 0.1$	$0.8 \pm 0.1$	$7.2 \pm 1.3$

<sup>a</sup> The precise  $K_m$  value cannot be reliably determined.

**Table 3**

Primers designated for Lpn1 amplification and cloning into individual expression vectors.

Vector	Forward primer (5'–3')	Reverse primer (5'–3')
pOPINE	AGGAGATATACCATGTGGAACGCGATAGGGCACC	GTGATGGTGATGTTTAATAAGTTTATTTTACCTTCGGC
pOPINE <sup>a</sup>	AGGAGATATACCATGGTCCCCTTACCTGAGTTAAGAAT G	GTGATGGTGATGTTTAATAAGTTTATTTTACCTTCGGCAA
pOPINMaIE	CCGCCTCGGCTCTCGCCTGGAACGCGATAGGGCACC	GTGATGGTGATGTTTAATAAGTTTATTTTACCTTCGGCAA
pOPINDsba	TTTAGCGCATCGGCGTGGAACGCGATAGGGCACC	GTGATGGTGATGTTTAATAAGTTTATTTTACCTTCGGCAA
pOPINO	CTACCGTAGCGCAAGCTTGGAACGCGATAGGGCACC	GTGATGGTGATGTTTAATAAGTTTATTTTACCTTCGGCAA
pOPINP	GCCCAGCCGGCGATGGCCATGTGGAACGCGATAGGGC ACC	GTGATGGTGATGTTTAATAAGTTTATTTTACCTTCGGCAA
pOPINTolB	CATCAGTTCTGCATGCTTGGAACGCGATAGGGCACC	GTGATGGTGATGTTTAATAAGTTTATTTTACCTTCGGCAA
pOPINS	GCGAACAGATCGGTGGTTGGAACGCGATAGGGCACC	ATGGTCTAGAAAGCTTTAAATAAGTTTATTTTACCTTCGG CAA

<sup>a</sup> Construct with native signal sequence

**Table 4**

Reaction buffers used for determination of pH optima of Lpn1.

Composition of reaction buffers	pH
50 mM Citric acid, 50 mM NaCl	3.0
50 mM Sodium acetate, 50 mM NaCl	4.0
50 mM Bis-Tris, 50 mM NaCl	5.0
50 mM Bis-Tris, 50 mM NaCl	5.5
50 mM Bis-Tris, 50 mM NaCl	6.0
50 mM Bis-Tris, 50 mM NaCl	6.5
50 mM Bis-Tris, 50 mM NaCl	7.0
50 mM Tris-HCl, 50 mM NaCl	8.0
50 mM Glycine, 50 mM NaCl	9.0

## Captions to illustrations

**Fig. 1.** Sequence alignment of Lpn1 and S1–P1 type nucleases from selected pathogens. Sequences from the following organisms are compared: *Legionella pneumophila* (Lpn1), *Legionella (Fluoribacter) bozemanii* (Fbz), WP\_082642378, *Legionella longbeachae* (Llon), WP\_003631713, *Pseudomonas aeruginosa* (Paer), CRR28597, *Stenotrophomonas maltophilia* (Smal), WP\_065175752, *Trypanosoma brucei gambiense* (Tbru), XP\_011774075, *Leishmania major* Friedlin, XP\_001684744, and *Plasmodium falciparum* MaliPS096\_E11 (Pfal), ETW46982. Signal peptides are included and marked. The mature enzyme N-terminal residue is the first consensus tryptophan residue. Amino acid residues involved in the formation of the zinc cluster (zinc ions Zn1, Zn2, and Zn3) and of NBS1 are marked. The figure was created using Clustal Omega [23] and ESPript 3.0 (<http://esprict.ibcp.fr>) [32].

**Fig. 2.** Denaturing gel electrophoresis and Western blot demonstrating Lpn1 expression in *E. coli* Lemo21 cells and purity after Zn-NTA and heparin affinity chromatography. (A) SDS-PAGE and (B) Identity of recombinant Lpn1 verified by Western blot using anti His-tag antibody: M – Color Prestained Protein Standard, lane 1 – non-induced bacterial cells carrying the pOPINMalE-Lpn1 construct as negative control, lane 2 – bacterial cells lysate from overnight Lpn1 production in Overnight Express<sup>TM</sup> Instant TB Medium, lane 3 – insoluble fraction of the whole cell bacterial lysate, lane 4 – soluble fraction of the whole cell bacterial lysate, lane 5 – Lpn1 after Zn-NTA, lane 6 – Lpn1 purified by heparin affinity chromatography.

**Fig. 3.** Kinetic profiles of Lpn1. Substrate inhibition is apparent for all used substrate types. Lpn1 activity is expressed in Units (U); 1 Unit of enzyme activity is defined as a change of absorbance (A260) of 0.001 in 1 cm path per 1 minute per 1 µg of Lpn1. Standard deviations are given from three measurements. Kinetic parameters are listed in Table 2. (A) Activity towards RNA and ssDNA substrates. (B) Activity towards dsDNA. (C) 3′ nucleotidase activity (3′-AMP as substrate). Enzyme activity is given as µmol of released phosphate per minute per 1 µg of Lpn1.

**Fig. 4.** pH optima and temperature optima of the catalytic activity of Lpn1. Standard deviations from three measurements are given. (A) The effect of pH on activity towards RNA and ssDNA. The optimum of activity towards RNA was determined around pH 7.0, while the maximal enzymatic activity towards ssDNA was observed at pH 6.0. The unit of enzyme activity was defined as a change of absorbance (A260) of 0.001 in 1 cm path per 1 minute per 1 µg of Lpn1. (B) pH dependence of Lpn1 nucleotidase activity. The highest nucleotidase activity was observed between pH values 6.0 and 7.0. The unit of enzyme activity was defined as µmol of released phosphate per minute per 1 µg of Lpn1. (C) Temperature dependence of RNase activity of Lpn1. The temperature optimum of RNA cleavage was observed around 65 °C. The unit of enzyme activity was defined as change of absorbance (A260) of 0.001 in 1 cm path per 1 minute per 1 µg of Lpn1.

**Fig. 5.** Electrostatic potential distribution demonstrated on a model of Lpn1 based on S1 nuclease compared to S1 nuclease (PDB id 5FBA). Electrostatic isosurface at levels –1 kT (red) and +1 kT (blue) is shown above the molecular surface (wheat) viewed from the side of the active site. (A) Electrostatics calculated for Lpn1 model at pH 6.0 and (B) at pH 7.0, and (C) for S1 nuclease at pH 4.0 (activity optimum). The graphics were prepared using PyMOL (Schrödinger, LLC).

**Fig. 6.** Alignment of Lpn1 sequence with all the other members of the S1–P1 family with determined three-dimensional structure. Sequence identity is marked by black background, high similarity within the group by grey background; signal peptides are excluded from the alignment.

Residues involved in zinc ion binding and formation of NBS1 are marked, as well as the region defining specificity. Residues in this region forming the positive surface patches in TBN1, the Tyr site in P1, and the Half-Tyr site in S1 are underlined. The alignment was prepared using Clustal Omega [23] and GeneDoc 2.7 [24].

**Fig. 7.** Representation of sequence homology to the most identical available three-dimensional structure of S1–P1 type nuclease TBN1 from *Solanum lycopersicum*. The surface of TBN1 (PDB id 3SNG) is colored according to sequence homology of Lpn1 to TBN1: red – identical residues, green – similar residues, dark blue – dissimilar residues, deletions in Lpn1 sequence are depicted as grey sticks in TBN1. The zinc ions of TBN1 are shown as cyan spheres. The key elements of the enzyme are maintained (zinc cluster, NBS1), while the specificity defining surface differs. The graphics was prepared using PyMOL (Schrödinger, LLC).

**Fig. 8.** Solution state and temperature stability of Lpn1 studied by circular dichroism. (A) CD spectrum of Lpn1 at room temperature. The protein is mainly  $\alpha$ -helical. (B) The melting temperature of Lpn1 determined by the change of the CD signal at 222 nm is 73 °C.

**Fig. 9.** Size exclusion chromatography profile of Lpn1 recorded at 280 nm using a Superdex 75 10/300 GL column (GE Healthcare).



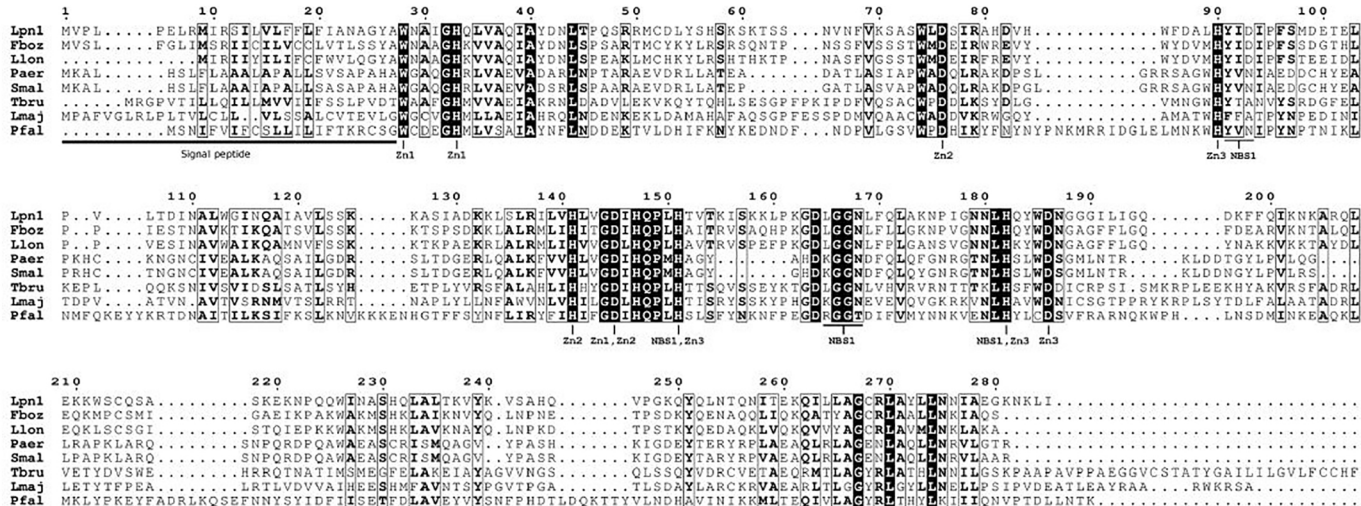


Figure 1

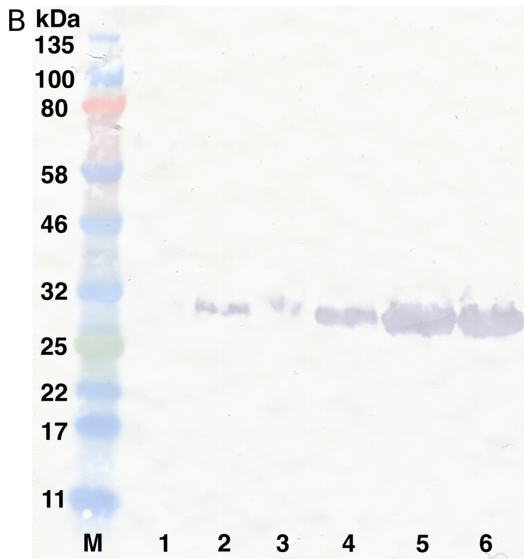
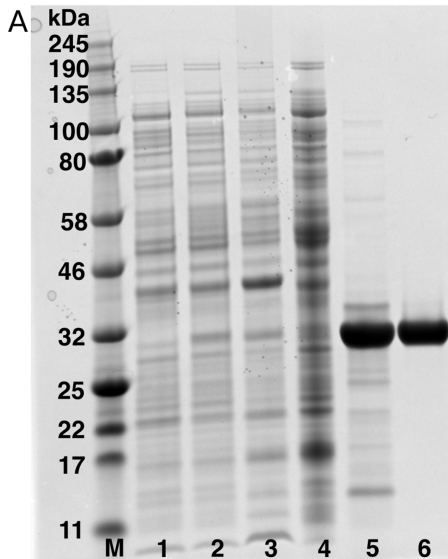


Figure 2

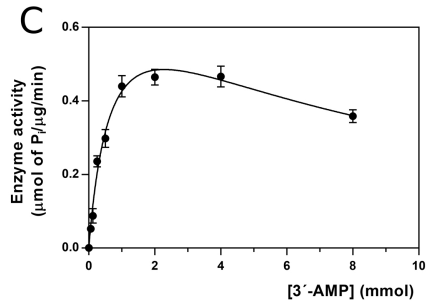
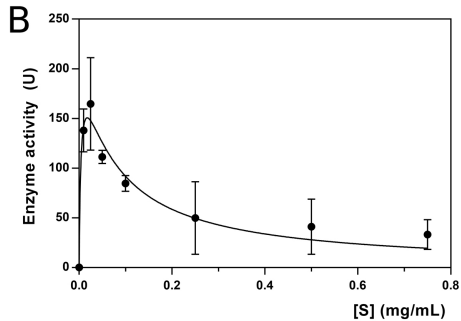
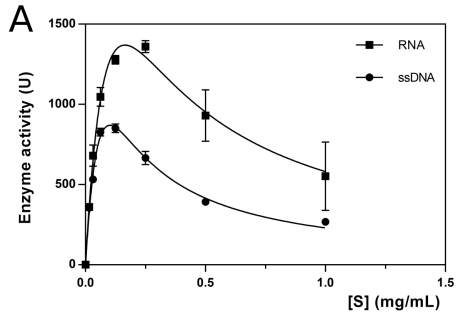


Figure 3

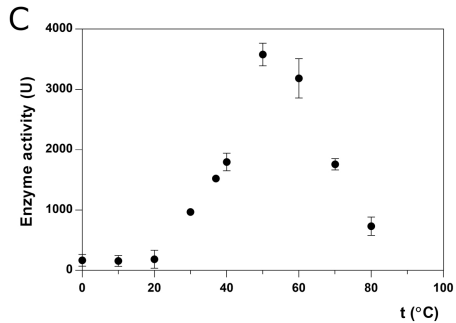
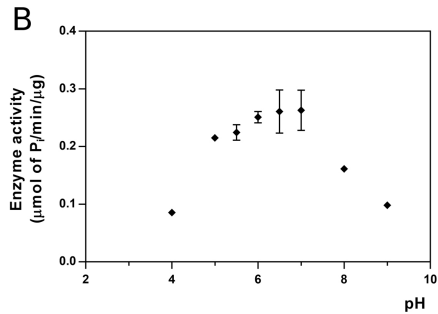
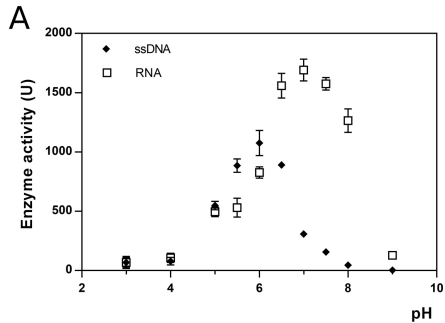


Figure 4

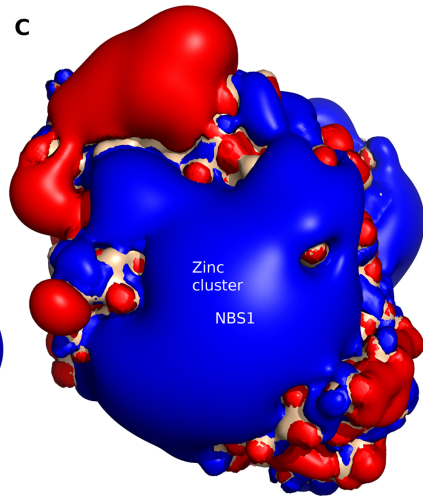
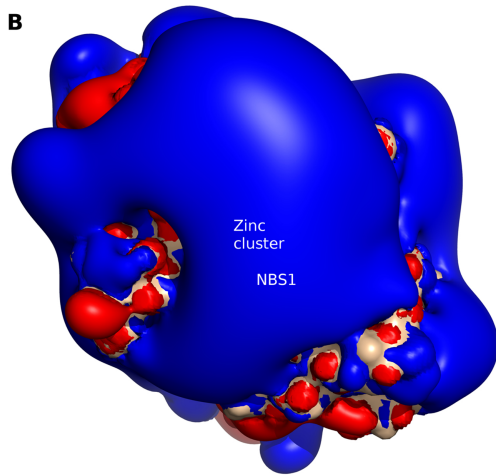
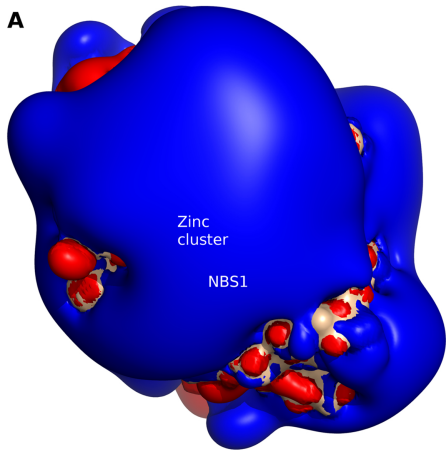


Figure 5

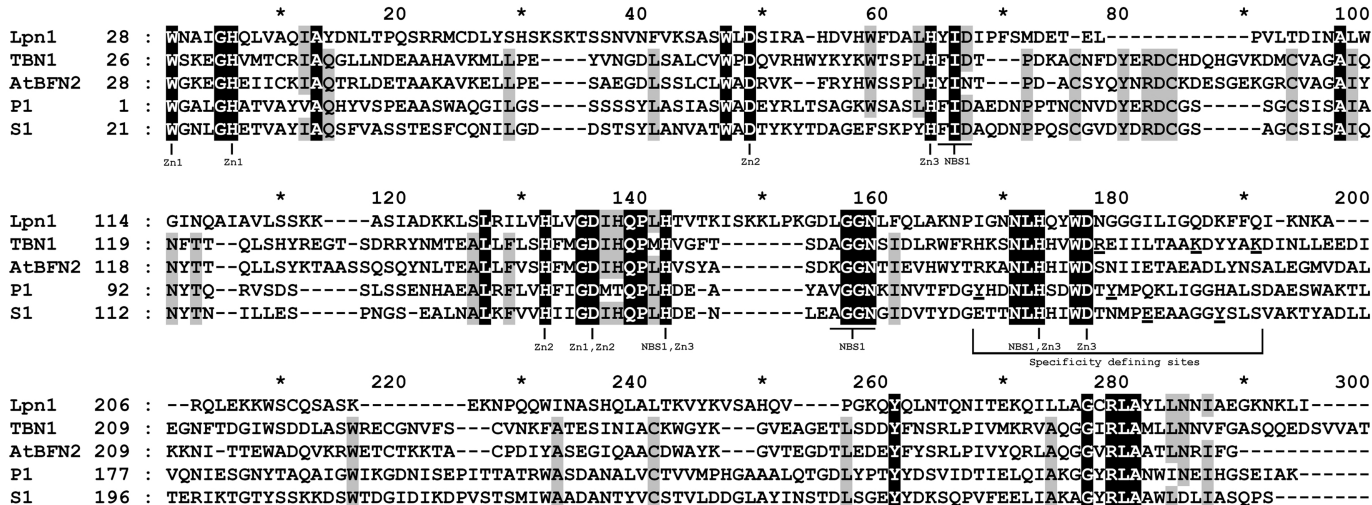


Figure 6

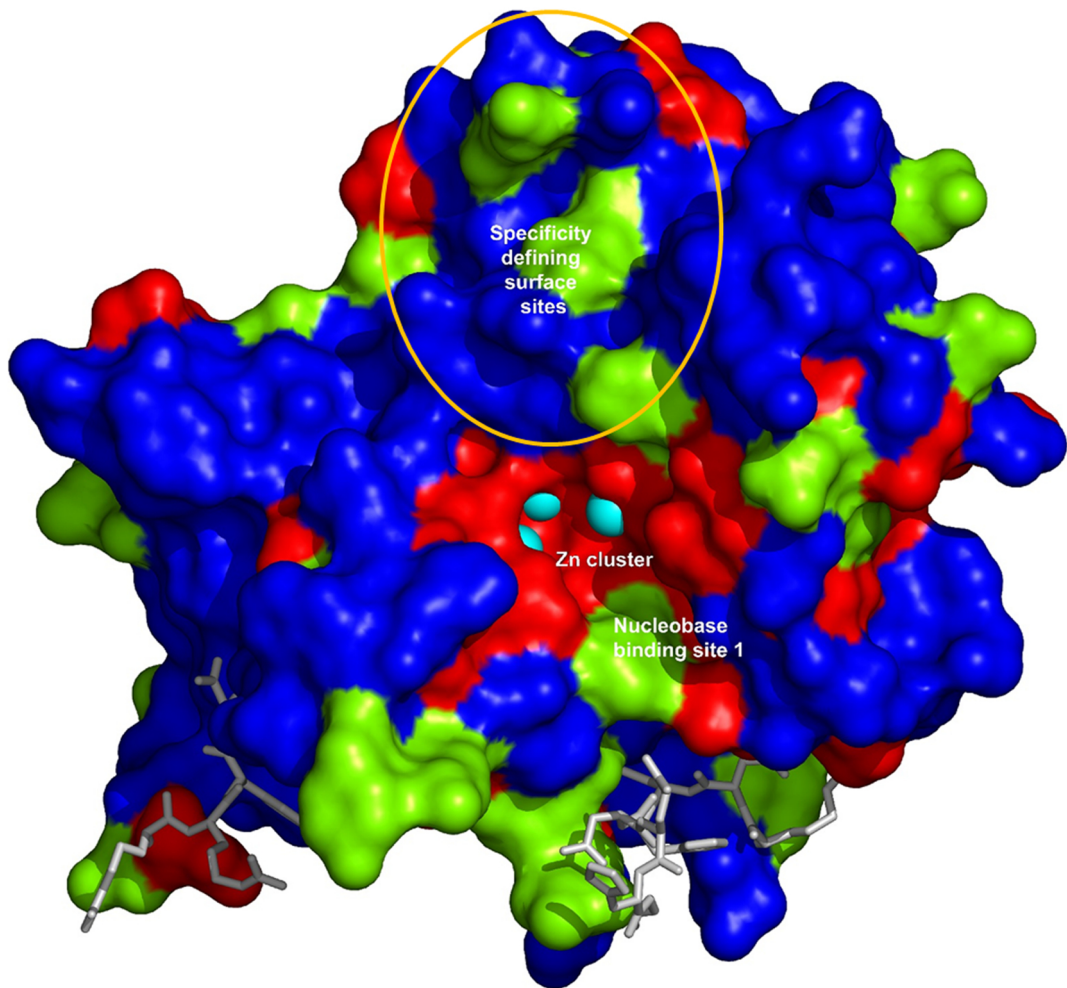


Figure 7

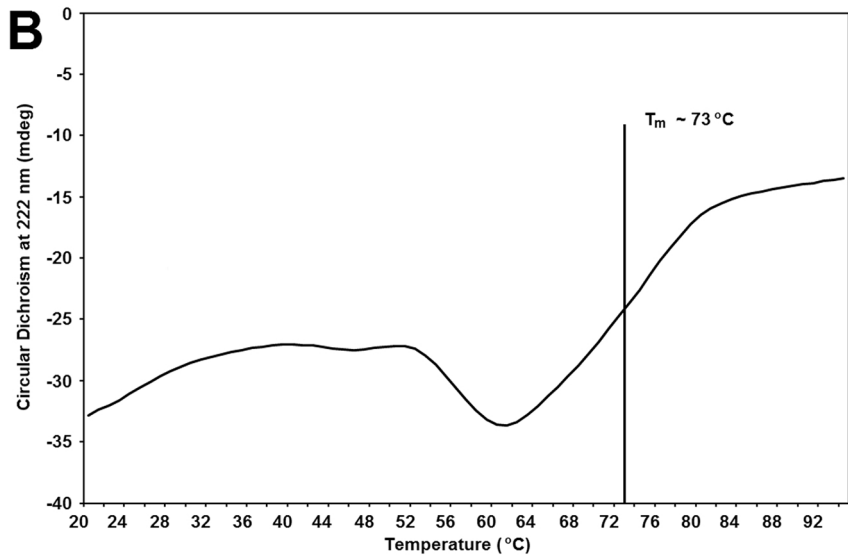
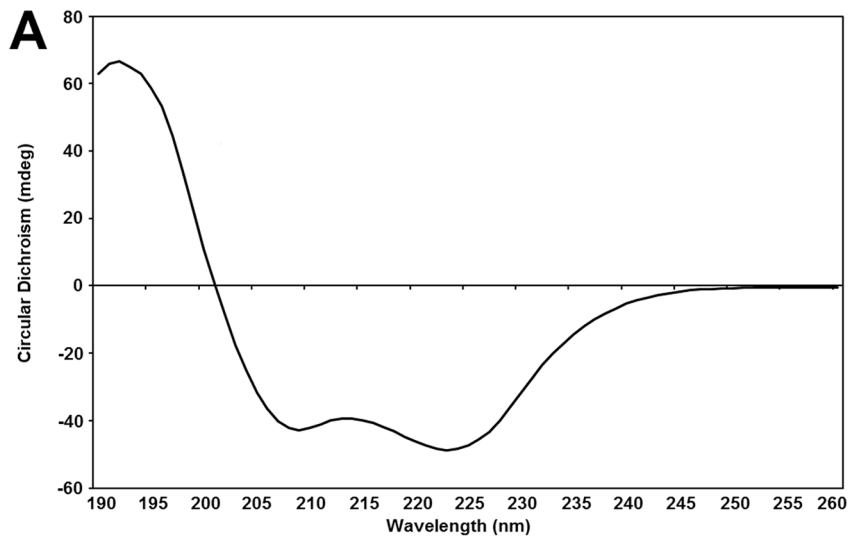


Figure 8



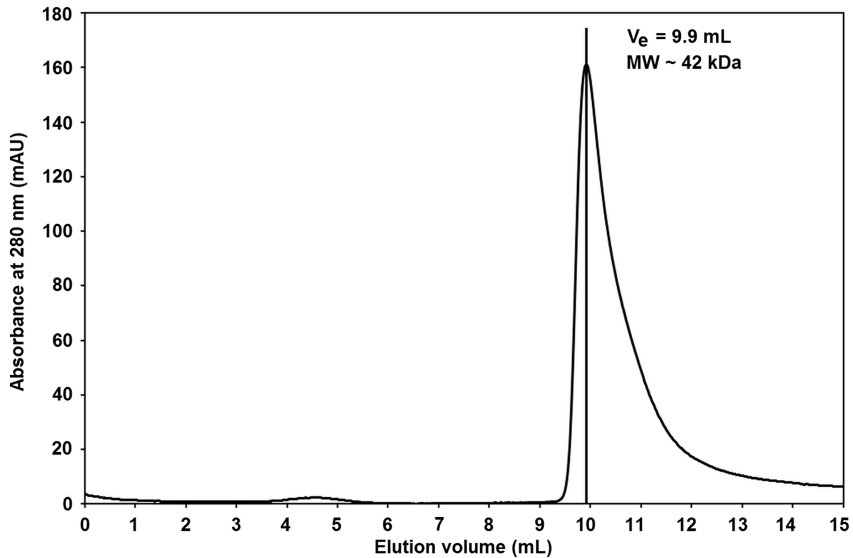


Figure 9