THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 287, NO. 24, pp. 20221–20230, June 8, 2012

Highly Sensitive Quenched Fluorescent Substrate of *Legionella* Major Secretory Protein (Msp) Based on Its Structural Analysis*[S]

Received for publication, December 22, 2011, and in revised form, April 12, 2012 Published, JBC Papers in Press, April 23, 2012, DOI 10.1074/jbc.M111.334334

Hervé Poras[‡1], Sophie Duquesnoy[‡], Emilie Dange[‡], Anthony Pinon[§], Michèle Vialette[§], Marie-Claude Fournié-Zaluski[‡], and Tanja Ouimet[‡]

From [‡]Pharmaleads, Paris BioPark, 11 Rue Watt 75013 Paris and the [§]Institut Pasteur Lille, Unité de Sécurité Microbiologique, 1 Rue du Professeur Calmette, 59019 Lille Cedex, France

Background: *Legionella pneumophila* secretes a protease without known specific substrate and three-dimensional structure.
Results: Analysis of a quenched peptide library identified a lead substrate, ameliorated by rational design, using an Msp structural model obtained by the x-ray structure of pseudolysin.
Conclusion: The study identifies the first selective substrate for Msp.
Significance: This substrate could be useful for *Legionella* detection.

Legionella pneumophila **has been shown to secrete a protease termed major secretory protein (Msp). This protease belongs to the M4 family of metalloproteases and shares 62.9% sequence similarity with pseudolysin (EC 3.4.24.26). With the aim of developing a specific enzymatic assay for the detection and quantification of Msp, the Fluofast substrate library was screened using both enzymes in parallel. Moreover, based on the crystal structure of pseudolysin, a model of the Msp structure was built. Screening of the peptide library identified a lead substrate specifically cleaved by Msp that was subsequently optimized by rational design. The proposed model for Msp is consistent with the enzymatic characteristics of the studied peptide substrates and provides new structural information useful for the characterization of the protease. This study leads to the identification of the first selective and high affinity substrate for Msp that is able to detect picomolar concentrations of the purified enzyme. The identified substrate could be useful for the development of a novel method for the rapid detection of *** **Legionella** **.**

Legionella is a genus of Gram-negative bacteria, widely distributed in freshwater environments, where it survives as intracellular parasites of protozoa. Among the different species of *Legionella*, the most frequently associated with human disease is *Legionella pneumophila*, causing acute pneumonia, referred to as legionellosis, and Pontiac fever, a severe influenza-like illness. In the human lung, the bacteria are able to replicate within alveolar macrophages and epithelial cells, causing tissue damage. Several potential virulence factors have been identified in *L. pneumophila*, including a peptide toxin, various enzymes (phosphatase, phospholipase C, proteases, and kinases), and the Mip gene product (1).

Among the proteases, one exotoxin, abundantly secreted by different strains of *L. pneumophila*, was detected (2) and shown to possess a proteolytic activity toward some human serum proteins (3), collagen, casein, gelatin, and hide powder but not elastin (4). This extracellular protease, designated as major secretory protein (Msp),[2] has been purified from *L. pneumophila* culture filtrates (5) and characterized as a neutral zinc-containing metalloprotease of 38 kDa, exhibiting hemolytic and cytotoxic activities (6).

The protease has been cloned (7), and analysis of the gene has shown that it encodes a large preproenzyme of 543 amino acids (60,775 Da) transported to the periplasma before maturation into the 38-kDa active protein. The protease belongs to the M4 family of metalloproteases (clan MA/E) represented by its archetypal member thermolysin. Within the M4 family, Msp shares the highest structural and functional homologies with pseudolysin (*Pseudomonas aeruginosa* elastase, EC 3.4.24.26) (8, 9). Comparison of their amino acid sequences allowed the characterization of the main residues involved in the proteolytic activity of Msp, *i.e.* the consensus sequences ^{377}HEVSH^{381} and ^{401}ESFSD^{405} containing the three zinc ligands His^{377}, His^{381}, Glu^{401}, and the catalytic glutamate Glu^{378}. As expected, the mutation of E378Q leads to an inactive protease deprived of cytotoxic activity (9), thus probing the involvement of the hydrolytic properties of the protease in *L. pneumophila* pathogenesis.

A chromogenic substrate MeO-succinimide-Arg-Pro-Tyr-*p*-nitroanilide (S-2586), initially developed for α-chymotrypsin (10), has been used to detect Msp in different strains of *Legionella* (11, 12). Despite the lack of specificity of this substrate, species-discriminating responses were obtained. All the sero-

* This work was supported by a grant from Agence Nationale pour la Recherche (Programme de Recherche sur les ECOtechnologies et le Développement Durable, Ralf project).
[S] This article contains supplemental Fig. 1.
[1] To whom correspondence should be addressed. Tel.: 33144066095; Fax: 33144066099; E-mail: herve.poras@pharmaleads.com.

[2] The abbreviations used are: Msp, major secretory protein; HPI, *N*-(1-carboxy-3-phenylpropyl)-phenylalanyl-α-asparagine; Nop, (*p*-nitro)-L-phenylalanine; Pya, L-pyrenylalanine; Fmoc, *N*-(9-fluorenyl)methoxycarbonyl; Orn, ornithine; hSer, homo-serine; hArg, homo-arginine.

groups of *L. pneumophila* cleaved this substrate, whereas other *Legionella* species (2 of 5), *Pseudomonas* strains (4 of 8), or Enterobacteriaceae strains (0 of 11) were cleaved poorly or were not responsive. These results suggested a relatively high specificity of *L. pneumophila* Msp as compared with analogous secreted activities of other bacteria. Msp has interesting properties, *i.e.* its abundant secretion, its possible implication in *L. pneumophila* pathogenicity, and its specificity toward other pathogens. Taking into account these properties, the aim of this study was to develop a highly sensitive and selective substrate of Msp as a marker of *L. pneumophila.*

Consequently, based on the crystal structure of pseudolysin, obtained with 1.5-Å resolution (13), a three-dimensional model of Msp was constructed and refined. The secreted Msp protease was purified from culture media of *L. pneumophila* Philadelphia 1 (CIP 103854T-ATCC 33152). The Fluofast library of intramolecularly quenched fluorescent substrates containing the Pya/Nop fluorophore/quencher pair (14, 15) was screened in parallel with purified Msp and pseudolysin. Results of this screening provided a lead selective substrate. Docking of this substrate in the three-dimensional model of Msp and sequence refinement based on the observed substrate-enzyme interactions allowed the synthesis of highly efficient and selective substrates of *L. pneumophila* Msp.

EXPERIMENTAL PROCEDURES

Reagents—Fmoc-protected amino acids, piperidine, *N*-methylpyrrolidone, dichloromethane, dicyclohexylcarbodiimide, 1-hydroxybenzotriazole, and *O*-(7-azabenzotriazol-1-yl)-1,1,3,3-tetramethyluronium were purchased from Applied Biosystems (Courtaboeuf, France). Fmoc-(*p*-Nitro)-L-phenylalanine and methylbenzhydrylamine resin (0.73 mmol/g) were from Novabiochem. Fmoc-L-pyrenylalanine (Pya), prepared as described previously (16), were from Polypeptide Laboratories (Strasbourg, France). Trifluoroacetic acid was from SDS-Carlo Erba (France). Triisopropylsilane was from Sigma. BCYEα (agar-solidified Buffered Charcoal Yeast Extract supplemented with 0.1% α-ketoglutarate) was obtained from Oxoid (Basingstoke, UK). Yeast extract was from Oxoid; L-cysteine HCl was from Merck, and ferric pyrophosphate from Sigma.

Pseudolysin (EC 3.4.24.26) was purchased from Calbiochem. Thermolysin (EC 3.4.24.27) was purchased from R&D Systems (France).

Culture—Msp was produced from *L. pneumophila* Philadelphia 1 strain (CIP 103854 T or ATCC 33152) using the protocol described by Ristroph *et al.* (17) with minor modifications. Briefly, stock cultures were maintained at −80 °C in a cryobank (AES Laboratoires, Combourg, France). The reference plating medium BCYEα (agar-solidified Buffered Charcoal Yeast Extract supplemented with 0.1% α-ketoglutarate) was used for counting (AFNOR NT T90-431).

The liquid growth medium was yeast extract broth, which was prepared as follows: yeast extract, 10 g/liter; L-cysteine HCl, 0.4 g/liter; ferric pyrophosphate, 250 mg/liter; and sterilized by filtration (0.2-μm membranes, Millipore, Billerica, MA). The pH was adjusted to 6.9 by addition of 1 N NaOH.

After resuscitation in 9 ml of yeast extract broth for 4 days at 36 °C, strains were subcultured at 1:30 for 3 days at 36 °C with continuous shaking. The final volume of each culture was 3 liters. A sample was removed for enumeration on BCYEα, and the remaining culture medium was centrifuged at 5000 rpm for 10 min at 20 °C to remove suspended matter. The supernatant was filter-sterilized on a 0.2-μm membrane (Millipore).

Purification—Msp was purified based on the protocol from Dreyfus and Iglewski (5). Proteins from the sterilized culture media were precipitated overnight at 4 °C under constant stirring with 65% ammonium sulfate. The media were then centrifuged at 14,000 rpm for 30 min at 4 °C, and the protein pellets were resuspended using 25 mM Tris-HCl, pH 7.8, 25 mM NaCl, 0.01% Triton X-100 (buffer A), loaded on a HiPrep 26/10 desalting column (GE Healthcare, Akta Systems), and eluted in the same buffer at a flow rate of 10 ml/min. The fractions containing Msp activity (identified by SDS-PAGE or enzymatic activity) were pooled (120 ml) and concentrated (~10 ml at 4 °C using Amicon Ultra-15 (cutoff of 10 kDa) (Millipore)). The concentrated fraction was loaded onto a HiPrep DEAE FF 16/10 column (GE Healthcare, Akta Systems) equilibrated with buffer A, and the protein was eluted according to a multistep gradient buffer B (25 mM Tris-HCl, pH 7.8, 1 M NaCl, 0.01% Triton X-100). The first step consisted of 15% of buffer B (6 column volumes), the second of 60% of buffer B (6 column volumes), and the last elution step was set at 100% of the same (5 column volumes). The fractions containing Msp activity, eluting at 60% buffer B, were pooled and concentrated to a volume of 200 μl at 4 °C. This concentrated fraction was loaded onto a HiLoad 16/60 Superdex 75 preparation grade (GE Healthcare, Akta Systems) and eluted in 500-μl fractions with 25 mM Tris-HCl, pH 7.2, 0.15 M NaCl, 0.01% Triton X-100 at a flow rate of 0.25 ml/min across 1.5 bed volumes. The pure fractions containing Msp activity were pooled and concentrated using Amicon Ultra 10-kDa ultrafiltration. Proteins were quantified using BCA (Pierce).

The purified preparation was electrophoresed on a 12% acrylamide denaturing gel, silver-stained, and quantified by densitometry using a BSA scale and the Quantity One program of Bio-Rad. Purity of the preparation was also evaluated using gel filtration on a high resolution Superdex 75 10/300 column. For this purpose, purified Msp (50 μl) was loaded on a pre-equilibrated column (50 mM Tris-HCl, pH 7.2, 0.15 M NaCl, 0.01% Triton X-100) and separated at a flow rate 0.2 ml/min.

Homology Modeling and Substrate Docking—L. pneumophila endopeptidase Msp model was generated based on the structure of the *P. aeruginosa* elastase complexed with the HPI inhibitor (Protein Data Bank code 1U4G). Sequences were aligned using ClustalW (18) and five three-dimensional models were generated with MODELLER Version 8 (19). The best model was selected based on MODELLER energy scoring function.

The three-dimensional coordinates of the substrates were generated with DiscoveryStudio Version 2.0 (20). Gold Version 4.2 (21) was used to perform flexible docking of the substrates in the binding sub-sites, which constitute the proposed active site of Msp. First, the tri-peptide Ac-Pya-Nop-Gly, selected as scaffold, was docked in the proposed model of Msp. For the longer substrates, scaffold constraints were extracted from the position of the former tripeptide and were imposed during

TABLE 1
List of synthetic fluorescently quenched substrates
hArg is homo-Arg; hSer is homo-Ser.

Compound **1**	Ac-SKG-Pya-Nop-Gly-Gly-Lys-NH_2
Compound **2**	Ac-Ser-Lys-Gly-Pya-(3-NO_2)Tyr-Gly-Gly-Lys-NH_2
Compound **3**	Ac-Ser-Lys-Gly-Pya-(3,5- $(NO_2)_2$)Tyr-Gly-Gly-Lys-NH_2
Compound **4**	Ac-Ser-Lys-Gly-Pya-(3-NO_2)Tyr-Gly-Lys-NH_2
Compound **5**	Ac-Ser-Lys-Gly-Pya-(3-NO_2)Tyr-Lys-NH_2
Compound **6**	Ac-Ser-Lys-Gly-Pya-(3-NO_2)Tyr-Orn-NH_2
Compound **7**	Ac-Ser-Lys-Gly-Pya-(3-NO_2)Tyr-β-Lys-NH_2
Compound **8**	Ac-Ser-Lys-Gly-Pya-(3-NO_2)Tyr-Arg-NH_2
Compound **9**	Ac-Ser-Lys-Gly-Pya-(3-NO_2)Tyr-hArg-NH_2
Compound **10**	Ac-Ser-hArg-Gly-Pya-(3-NO_2)Tyr-Gly-Gly-Lys-NH_2
Compound **11**	Ac-Ser-Orn-Gly-Pya-(3-NO_2)Tyr-Gly-Gly-Lys-NH_2
Compound **12**	Ac-Ser-Lys-Glu-Pya-(3-NO_2)Tyr-Gly-Gly-Lys-NH_2
Compound **13**	Ac-Ser-Nle-Gly-Pya-(3-NO_2)Tyr-Gly-Gly-Lys-NH_2
Compound **14**	Ac-Arg-Gly-Pya-(3-NO_2)Tyr-Gly-Gly-Lys-NH_2
Compound **15**	Ac-Ser-Arg-G-Pya-(3-NO_2)Tyr-Gly-Gly-Lys-NH_2
Compound **16**	Ac-hSer-Lys-Gly-Pya-(3-NO_2)Tyr-Orn-NH_2
Compound **17**	Ac-Ser-Lys-Pya-(3-NO_2)Tyr-hArg-NH_2
Compound **18**	Ac-Lys-Pya-(3-NO_2)Tyr-hArg-NH_2
Compound **19**	Ac-Pya-(3-NO_2)Tyr-Orn-NH_2

the docking process. The constraints are applied at the placement stage by forcing all atoms on matching substructure of the ligand to match with the position of the corresponding scaffold atoms. The positions generated by Gold were graphically inspected with DSViewerPro, and a model compatible with identified SAR was selected to describe the substrate environment.

Synthesis—A quenched fluorescent peptide library corresponding to the general formula Ac-SKG-Pya-$(X)_n$-Nop-GGK-NH_2 has been developed for the characterization of all types of proteolytic activities (15, 16). In this formula, Pya (L-pyrenylalanyl) is the fluorescent moiety; Nop (L-4-NO_2-phenylalanyl) the quencher moiety; *X* indicates either a single natural amino acid chosen from Ala, Leu, Ile, Glu, Thr, Gln, Phe, Trp, Pro, or Lys or a mixture of these 10 amino acids, and *n* varies from 0 to 3. This library has been prepared by solid phase synthesis, using the split and mix method (22). When $n = 0$, the sub-library is constituted by one peptide; when $n = 1$, the sub-library is a batch of 10 peptides; when $n = 2$, the sub-library corresponds to ten batches of 10 peptides, and for $n = 3$ the sub-library corresponds to 100 batches of 10 peptides.

The individual peptides synthesized for the optimization of Msp recognition were also obtained by solid phase method using the Fmoc chemistry on methylbenzhydrylamine resin using an Applied Biosystems ABI 433 synthesizer, and a classical 2-(1*H*-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate/hydroxybenzotriazole protocol.

The syntheses proceeded in *N*-methylpyrrolidone with 10 eq of amino acids. The L-pyrenylalanine amino acid was introduced by mixing the resin in a syringe using (benzotriazol-1-yloxy)tris(dimethylamino)phosphonium hexafluorophosphate as coupling reagent in the presence of diisopropylethylamine. Deprotection of the lateral chains of the amino acids was accomplished by mixing the resin in presence of TFA, triisopropylsilane, and water (95:2.5:2.5) for 2 h.

Peptides were purified by semi-preparative HPLC using a Waters 600 device equipped with a 2487 dual absorbance detector and analyzed on an ACE C18, 100 Å, 5 μm, 250 × 4.6 mm column (unless otherwise specified) using CH_3CN (0.1% TFA)/H_2O (0.1% TFA) as elution system and a flow rate of 1 ml/min (Table 1). The following compounds were used: compound **1**, Ac-Ser-Lys-Gly-Pya-(4-NO_2)Phe-Gly-Gly-Lys-NH_2, HPLC CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 30–40% in 30 min; RT, 9.65 min; ESI (+): $[(M + 2H)/2]^+ = 519.4$. Compound **2**, Ac-Ser-Lys-Gly-Pya-(3-NO_2)Tyr-Gly-Gly-Lys-NH_2, HPLC CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 10–90% in 30 min; RT, 13.69 min; ESI (+): $[(M+2H)/2]^+ = 527.4$. Compound **3**, Ac-Ser-Lys-Gly-Pya-(3, 5-$(NO_2)_2$)Tyr-Gly-Gly-Lys-NH_2, HPLC CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 28/72; RT, 21.56 min; ESI (+): $[(M + 2H)/2]^+ = 550.0$. Compound **4,** Ac-Ser-Lys-Gly-Pya-(3-NO_2)Tyr-Gly-Lys-NH_2, HPLC CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 10–90% in 30 min; RT, 13.09 min; ESI (+): $[(M + 2H)/2]^+ = 498.9$. Compound **5**, Ac-Ser-Lys-Gly-Pya-(3-NO_2)Tyr-Lys-NH_2, HPLC CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 30/70; RT, 12.87 min. ESI (+): $[(M + 2H)/2]^+ = 470.3$. Compound **6**, Ac-Ser-Lys-Gly-Pya-(3-NO_2)Tyr-Orn-NH_2, HPLC Atlantis T3, 3.5 μm, 100 × 4.6 mm CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 10–90% in 15 min; RT, 9.02 min; ESI (+): $[(M+2H)/2]^+ = 477.8$. Compound **7**, Ac-Ser-Lys-Gly-Pya-(3-NO_2)Tyr-β-Lys-NH_2, HPLC Atlantis T3, 3.5 μm, 100 × 4.6 mm CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 10–90% in 15 min; RT, 8.97 min; ESI (+): $[(M + 2H)/2]^+ = 463.5$. Compound **8**, Ac-Ser-Lys-Gly-Pya-(3-NO_2)Tyr-Arg-NH_2, HPLC CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 30/70; RT, 10.28 min; ESI (+): $[(M + 2H)/2]^+ = 484.0$. Compound **9**, Ac-Ser-Lys-Gly-Pya-(3-NO_2)Tyr-hArg-NH_2, HPLC CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 30/70; RT, 11.25 min; ESI (+): $[(M + 2H)/2]^+ = 491.5$. Compound **10**, Ac-Ser-hArg-Gly-Pya-(3-NO_2)Tyr-Gly-Gly-Lys-NH_2, HPLC Atlantis T3, 3.5 μm, 100 × 4.6 mm CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 30/70; RT, 13.21 min; ESI (+): $[(M+2H)/2]^+ = 548.5$. Compound **11**, Ac-Ser-Orn-Gly-Pya-(3-NO_2)Tyr-Gly-Gly-Lys-NH_2, HPLC CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 25/75; RT, 41.62 min; ESI (+): $[(M + 2H)/2]^+ = 520$. Compound **12**, Ac-Ser-Lys-Glu-Pya-(3-NO_2)Tyr-Gly-Gly-Lys-NH_2, Atlantis T3, 3.5 μm, 100 × 4.6 mm CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 10–90% in 15 min; RT, 8.44 min; ESI (+): $[(M + 2H)/2]^+ = 563.4$. Compound **13**, Ac-Ser-Nle-Gly-Pya-(3-NO_2)Tyr-Gly-Gly-Lys-NH_2, HPLC CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 35/65; RT, 10.24 min; ESI (+): $[(M + 2H)/2]^+ = 519.9$. Compound **14**, Ac-Arg-Gly-Pya-(3-NO_2)Tyr-Gly-Gly-Lys-NH_2, HPLC Atlantis T3, 3.5 μm, 100 × 4.6 mm CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 28/72; RT, 15.40 min; ESI (+): $[(M + 2H)/2]^+ = 497.8$. Compound **15**, Ac-Ser-Arg-Gly-Pya-(3-NO_2)Tyr-Gly-Gly-Lys-NH_2, HPLC CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 0–30% in 20 min; then 30% 10 min; RT, 29.83 min; ESI (+): $[(M + 2H)/2]^+ = 541.4$. Compound **16**, Ac-hSer-Lys-Gly-Pya-(3-NO_2)Tyr-Orn-NH_2, HPLC CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 30/70; RT, 13.70 min; ESI (+): $[(M + 2H)/2]^+ = 470.5$. Compound **17**, Ac-Ser-Lys-Pya-(3-NO_2)Tyr-hArg-NH_2, HPLC CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 10–90% in 30 min; RT, 22.3 min. ESI (+): $[(M + 2H)/2]^+ = 462.8$. Compound **18**, Ac-Lys-Pya-(3-NO_2)Tyr-hArg-NH_2, HPLC CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 10–90% in 30 min; RT, 22.82 min; ESI (+): $[(M + 2H)/2]^+ = 419.2$. Compound **19**, Ac-Pya-(3-NO_2)Tyr-Orn-NH_2, HPLC CH_3CN (0.1% TFA)/H_2O (0.1% TFA) 50/50; RT, 4.18 min; ESI (+): $[(M + 2H)/2]^+ = 653.7$. Fluorescent metabolite: Ac-Ser-Lys-Gly-Pya, HPLC CH_3CN (0.1% TFA)/H_2O (0.1%

TFA) 10–90% in 30 min; RT, 14.2 min; ESI (+): $[(M + H)]^+ = 604.4$.

Enzymatic Assays—The substrate library was screened using 30 ng/ml purified Msp in 50 mM HEPES, pH 7.0, and pseudolysin in 50 mM Tris-HCl, pH 7.2, 2.5 mM $CaCl_2$. The sub-libraries ($n = 1, 2, 3$), which contain a mixture of 10 peptides, were tested at a final concentration of 100 μM, and the peptide corresponding to $n = 0$ at 10 μM. The enzymatic reactions were left to proceed at 37 °C for 180 min.

Lead substrate optimization studies were performed using purified substrates (10 μM) and 10 ng/ml of purified Msp. Specificity studies were also performed using thermolysin. For this purpose, substrates (10 μM) were incubated in parallel in reaction buffer containing 50 mM Tris-pH 7.5, 10 mM $CaCl_2$, 150 mM NaCl, and 0.05% Brij 35. All assays were performed in half 96-well low binding black plates in a final volume of 100 μl. Direct fluorescence monitoring of substrate hydrolysis was performed using a microplate fluorimeter (Berthold Series Twinkle^{TM} LB970) coupled to Mikrowin 2000 software with excitation at 340 nm, emission at 405 nm, lamp energy at 10,000, and temperature at 37 °C.

Kinetic Parameters—Determinations of K_m values were performed in initial rate conditions. Compounds **1**, **2**, **6**, and **9** at six different concentrations ranging from 1 to 30 μM were incubated with 2 ng/ml purified Msp in 50 mM HEPES, pH 7.0, for 60 min (compound **1**) or 30 min (compounds **2**, **6**, and **9**) at 37 °C. For each substrate concentration, a standard calibration was established for 5, 10, 15, and 20% cleavage by using the corresponding mixture of substrate and the fluorescent cleavage product (Ac-SKG-Pya). The obtained fluorescence was then reported to each standard calibration to quantify the amount of fluorescent cleavage product formed. The apparent K_m values (mean of three independent assays in duplicate) were calculated using the Graphpad Prism 4 program, and the k_{cat} was determined using the equation $k_{cat} = V_{max}/[E]$ with an enzyme molar mass of 37 kDa. Fluorescent signals were monitored using a Berthold Series Twinkle TM LB970 with excitation at 340 nm, emission at 405 nm, and lamp energy at 10,000.

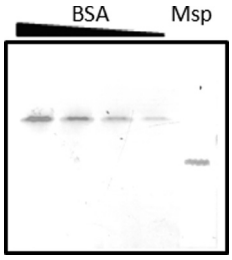

FIGURE 1. **SDS-PAGE analysis of purified Msp.** An Msp sample (2.5 μl) was electrophoresed in nonreducing conditions on a 12% polyacrylamide gel in parallel with BSA standards. The gel was silver-stained, and the Msp band was quantified by densitometric analysis against the BSA standards of 500, 250, 100, and 50 ng. Using this BSA scale, the Msp concentration is evaluated at 100 ng/μl.

RESULTS

Culture

Starting from the classical culture method described by Ristroph *et al.* (17), with some minor improvements allowing experiments to be performed in large volumes of growth medium, we obtained efficient and reproducible conditions of culture of *L. pneumophila* Philadelphia I. Several 3-liter batches of culture were obtained with bacteria concentrations in the 10^9–10^{11} cfu/liter range.

Purification

Starting from the filter-sterilized culture supernatants, about 1500 μg of pure enzyme were typically obtained as determined by gel filtration chromatography as well as by electrophoresis (Fig. 1 and Table 2). The pure Msp observed by gel electrophoresis displayed a molecular mass of 37 kDa and migrated as a single band.

Screening of the Fluofast Library Using Msp and Pseudolysin

The Ac-SKG-Pya-$(X)_n$-Nop-GGK-NH_2 sub-libraries were screened using both Msp and pseudolysin in parallel. Results obtained by screening the $n = 0$, 1, and 2 sub-libraries are depicted in Fig. 2. Very interestingly, the Ac-SKG-Pya-Nop-GGK-NH_2 peptide (compound **1,** $n = 0$), produced a very intense fluorescent signal in the presence of Msp, consistent with a cleavage between Pya and Nop. On the contrary, compound **1** was not cleaved by pseudolysin. Msp significantly cleaved only two $n = 2$ sub-libraries, -EX- and -Q*X*-, whereas pseudolysin efficiently cleaved all sub-libraries. Further screening of the $n = 3$ sub-libraries showed that they were all cleaved by pseudolysin (data not shown).

This first result suggests a difference in the structure of the S_1' subsites (nomenclature of Schechter and Berger (23)) of the two enzymes, with Msp able to accommodate the bulky aromatic Nop residue in its S_1' subsite but not pseudolysin. Conversely, the sub-libraries containing one or more residues between the fluorophore-quencher pair were well recognized by pseudolysin and poorly, or not at all, by Msp.

To further explore this difference, the cleavage products obtained after incubation of the $n = 1$ sub-library (Ac-SKG-Pya-*X*-Nop-GGK-NH_2) with either pseudolysin or Msp were analyzed by LC/MS. With either enzyme, only one fluorescent metabolite was observed, Ac-SKG-Pya, in proportions that were in accordance with the difference in the fluorescent signal previously observed (Fig. 2). With pseudolysin, three nonfluorescent metabolites, *X*-Nop-GGK-NH_2, with *X* = Phe, Trp, and

TABLE 2
Purification of Msp from *L. pneumophila* culture supernatant
AU means arbitrary unit.

Material	Volume	Protein	Total protein	Total activity	Recovery	Specific activity
	ml	*μg/ml*	*mg*	*AU*	*%*	*AU/μg*
Culture supernatant	3100	3384	10,490	1.4×10^8		13.35
Desalted precipitate	22	975	21.44	8.16×10^7	58	3806
HiPrep DEAE FF 16/10	120	42	5.05	1.1×10^8	134	21,825
HiLoad 16/60 Superdex 75	20	75	1.50	6.0×10^7	54	40,000

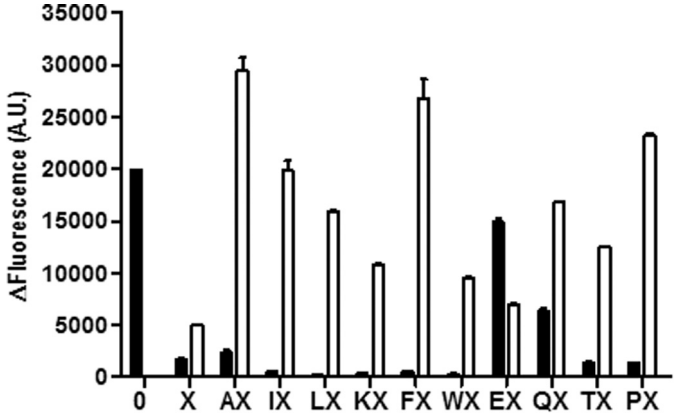

FIGURE 2. **Comparative screening of Ac-SKG-Pya-X*n*$_{(0-2)}$-Nop-GGK-CONH$_2$ peptide sub-libraries.** Either purified Msp (*black*) or pseudolysin (*white*) at 30 ng/ml was incubated at 37 °C for 3 h with 10 μM of compound **1** ($n = 0$) or 100 μM of each peptide sub-library in a final volume of 100 μl of 50 mM HEPES, pH 7, or 50 mM Tris-HCl, pH 7.2, 2.5 mM $CaCl_2$ reaction buffer, respectively. Control base fluorescence was measured by incubating the peptide library alone in 100 μl of its reaction buffer. The *histogram* shows the deltas of fluorescence corresponding to the fluorescence of the assay minus that of the base fluorescence of the library at 180 min (λ_{ex} = 340 nm and λ_{em} = 405 nm). *A.U.*, arbitrary units.

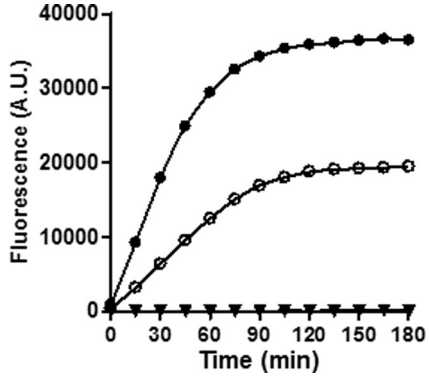

FIGURE 3. **Comparative cleavage of compounds 1–3 by Msp.** Purified Msp at 10 ng/ml was incubated at 37 °C for 180 min with either compounds **1** (○, Ac-SKG-Pya ↓ Nop-GGK-NH_2), **2** (●, Ac-SKG-Pya ↓ (3-NO_2)YGGK-NH_2), or **3** (▼, Ac-SKG-Pya-((3,5-$(NO_2)_2$)YGGK-NH_2) (10 μM) in a final volume of 100 μl of 50 mM HEPES, pH 7, and the fluorescent signal was read every 15 min (λ_{ex} = 340 nm and λ_{em} = 405 nm).

(Leu/Ile) were detected (data not shown), whereas Msp produced only the X = Phe- and Trp-containing nonfluorescent metabolites. This result further confirms the differences in the S_1' subsites of pseudolysin and Msp, but it also suggests that the ideal position of the Pya moiety lies in the S_1 subsite for both enzymes. With Pya in this position, the ensemble of the results shows that although Nop binds within the S_1' subsite of Msp, it is not accepted in the subsite of pseudolysin. The results of this screening revealed that a single peptide from the entire library (Ac-SKG-Pya-Nop-GGK-NH_2 (compound **1**) possessed the characteristics of a selective substrate of Msp.

Optimization of the Fluorigenic Specific Substrate of Msp

To further increase the performances of this selective substrate, rational modifications were introduced in the sequence of compound 1.

Replacement of Quencher Nop Moiety by 3-NO_2-Tyr and 3,5-Dinitro-Tyr—Comparison of the fluorescence increase obtained during the incubation of the three peptides compound **1** (Nop), compound **2** (3-NO_2-Tyr), and compound **3** (3, 5-dinitro-Tyr) with the same quantity of Msp in the same experimental conditions is presented in Fig. 3. These results show that although compound **3** was not cleaved, the fluorescent signal emitted as a result of the cleavage of compound **2** was about 2-fold greater than that obtained using compound **1**.

Modifications of Peptide Sequence Surrounding the Fluorophore/Quencher Pair—Modifications were first introduced in the C-terminal part of compound **2**, which was shortened by the sequential suppression of the two glycine residues (Fig. 4*A*, Table 1, and compounds **4** and **5**). Both peptides were substrates of Msp with the latter more rapidly cleaved than the parent compound **2**. However an external quenching, most probably due to the nonfluorescent metabolite, was responsible for a lower fluorescence increase.

Replacement of the C-terminal lysine residue in compound **5** by other positively charged residues (compounds **6** (Orn), **7** (β-Lys), **8** (Arg), and **9** (hArg)) (Fig. 4*A*) brings to light the importance of the charged side chain length and position for substrate efficiency. Compound **6** (Orn) was cleaved slightly more rapidly than compound **2**, but the increase in the fluorescent signal was lower (about 10%). Cleavage of compound **9** (hArg) by Msp also resulted in a good fluorescent signal, although this substrate was cleaved more slowly than compound **2** (Fig. 4*A*), and compounds **7** and **8** were not efficiently cleaved (data not shown).

Modifications were also introduced in the N-terminal domain of compound **2**. Results obtained using these substrates show that shortening the substrate N-terminal end produced a very significant decrease in the measured fluorescent signal. Although modification of the nature of the charged residue in compounds **10**, **11**, and **12** did not produce better substrates (data not shown), its absolute necessity was underlined by the lack of cleavage observed using compound **13** in which the lysine residue in the P_3 position was replaced by norleucine (Fig. 4*B*). Finally, combined N- and C-terminal modifications (compounds **14**-**19**) did not improve the fluorescent signal emitted after Msp cleavage, although compounds **17** and **18**, the two shortest substrates, were efficiently cleaved by Msp (Fig. 4*C*).

Selectivity of Compounds 2, 6, and 9 toward Msp Versus Pseudolysin and Thermolysin

Three peptides emerged from this screening as highly efficient Msp substrates: compounds **2**, **6**, and **9** (Table 1 and Fig. 4*A*). It was important to verify that the chemical modifications introduced in the primary sequence of compound **1** did not alter their specificity *versus* other M4 family of metalloproteases, *i.e.* pseudolysin and thermolysin. In the same experimental conditions of time, temperature, and enzyme concentration (120 min at 37 °C using 10 ng/ml enzyme), a small cleavage of compounds **6** and **9** was observed (mean ΔF of 171 and 337 arbitrary units, respectively) in the presence of pseudolysin, although no fluorescence increase could be detected using compound **2** (data not shown). Moreover, although compounds **6** and **9** were slightly but significantly cleaved by thermolysin (mean ΔF = 275 and 844 arbitrary units, respectively, in 120 min), compound **2** in the same enzymatic

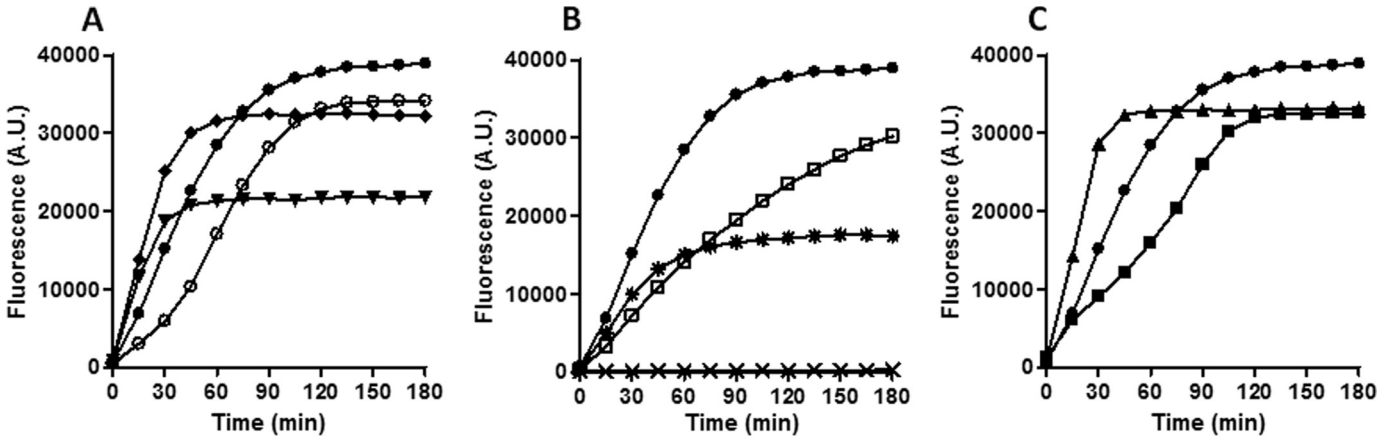

FIGURE 4. **Comparative cleavage of compounds 2 and 4–19 by Msp.** Purified Msp (10 ng/ml) was incubated at 37 °C for 180 min with compound **2** (●, Ac-SKG-Pya ↓ (3-NO_2)YGGK-NH_2) and either N- or C- or N- and C-terminal modified peptides (10 μM) in a final volume of 100 μl of 50 mM HEPES, pH 7, and the fluorescent signal was read every 15 min (λ_{ex} = 340 nm and λ_{em} = 405 nm). Compounds **5** (▼, Ac-SKG-Pya ↓ (3-NO_2)YK-NH_2), **6** (♦, Ac-SKG-Pya ↓ (3-NO_2)Y-Orn-NH_2), and **9** (○, Ac-SKG-Pya ↓ (3-NO_2)YhR-NH_2; where hR is homo-arginine) containing C-terminal modifications are compared with compound **2** cleavage in *A*, whereas fluorescent profiles obtained by the cleavage of compounds **12** (□, Ac-SKE-Pya ↓ (3-NO_2)YGGK-NH_2), **13** (×, Ac-S-Nle-G-Pya ↓ (3-NO_2)YGGK-NH_2), and **15** (*, Ac-SRG-Pya ↓ (3-NO_2)YGGK-NH_2) modified in their N terminus are represented in *B*. *C* shows the cleavage kinetics of compounds **17** (▲) and **18** (■) containing both C- and N-terminal modifications.

TABLE 3

Kinetic parameters of the key substrates Michaelis and Menten (K_m) and catalytic constants (k_{cat}) of the four best substrates were measured as described under "Experimental Procedures"

Results presented are the means of three independent assays (n = 3), each performed in duplicate. hR means homo-Arg.

Peptide	K_m	k_{cat}	k_{cat}/K_m
	μM	s^{-1}	$M^{-1}\cdot s^{-1}$
Ac-SKG-Pya ↓ Nop-GGK-NH_2, compound **1**	7.2 ± 1.7	12.1 ± 0.9	1.7×10^6 ± 0.2
Ac-SKG-Pya ↓ (3-NO_2)YGGK-NH_2, compound **2**	7.2 ± 1.5	25.7 ± 1.7	3.6×10^6 ± 0.2
Ac-SKG-Pya ↓ (3-NO_2)Y-Orn-NH_2, compound **6**	2.8 ± 0.4	14.5 ± 1.2	5.8×10^6 ± 0.7
Ac-SKG-Pya ↓ (3-NO_2)Y-hR-NH_2, compound **9**	1.1 ± 0.2	12.6 ± 0.6	12.0×10^6 ± 0.6

conditions gave only a background signal of 41 arbitrary units of fluorescence (data not shown).

Kinetic Parameters of Compounds 1, 2, 6, and 9

The kinetic parameters of the four best compounds were characterized (Table 3). Although the Michaelis and Menten constants (K_m) of compounds **1** and **2** were both found to be around 7 μM, modification of the C-terminal charged residue in compounds **6** and **9** resulted in a 2–3-fold increase in their K_m values (2.8 ± 0.4 and 1.1 ± 0.2, respectively), suggesting a better interaction of the substrate with the protein backbone. These improved interactions had a negative impact on the measured k_{cat} of the same compounds. Indeed, whereas compound **2** had a k_{cat} of almost 26 s^{-1}, those of compounds **6** and **9** were of 14.5 ± 1.2 and 12.6 ± 0.6 s^{-1}, respectively, values that were equivalent to that of compound **1** (Table 3).

Although the specificity constant (k_{cat}/K_m) of compound **9** toward Msp was the highest of the four studied substrates, with a $k_{cat}/K_m = 12.0 \times 10^6$ M^{-1} s^{-1}, and because the aim was to develop the most selective and sensitive substrate for the specific detection of *L. pneumophila,* via a quantification of the secreted Msp, compound **2** was preferred, for its complete selectivity and intense fluorescent signal was released upon cleavage by Msp.

Determination of the limit of detection and quantification of Msp using compound **2** (10 μM) reveals the capacity of the assay to detect as low as 6 pg/ml Msp (limit of detection = mean fluorescence of compound **2** + 3 S.D.) in 60 min. Standard curves using decreasing concentrations of Msp ranging from 0.06 to 0.006 ng/ml and calculation of 90% intervals allow the determining a limit of quantification of the assay of 0.02 ng/ml of Msp (Fig. 5).

Homology Modeling

The complete sequence of the preproprotease includes 543 amino acids distributed as follows: (i) a signal peptide (residues 1–24); (ii) a propeptide (residues 25–207), and (iii) the mature protease of interest (residues 208–543) (8).

Alignment of the pseudolysin and the target Msp sequences reveals a 47.7% sequence identity and 62.9% sequence similarity for the residues 212–543 of Msp (Fig. 6). The high sequence similarity together with the conserved catalytic residues (His, Glu, and His) strongly suggest that the pseudolysin structure constitutes a good template to build an accurate model of our target protease. This alignment indicates the presence of three insertion domains corresponding to the segments ^{239}RDSSV^{243}, ^{277}PDTQSTKTTYTGYSAD^{292}, and ^{329}KSDGSP^{334} (Fig. 7). In the absence of a corresponding template structure, these loops were generated *de novo* and are thus not optimized. In the final model, we verified that none of the insertion domain residues were in close proximity to the catalytic site.

The three-dimensional structure of the template pseudolysin was stabilized, not only by the catalytic Zn^{2+} ion but also by two disulfide bridges, one at the N terminus of the sequence (Cys^{30}–Cys^{58}) and the other at the C terminus (Cys^{270}–Cys^{297}), and by

a Ca^{2+} ion that is hexa-coordinated by five residues of the protein (Asp^{136}, Glu^{172}, Glu^{175}, Asp^{183}, and the CO of Leu^{185}) and a water molecule. Our proposed model is compatible with the conservation of the disulfide bridges observed in the template pseudolysin structure. The Msp N-terminal cysteine residues Cys^{246}–Cys^{273} are aligned with the corresponding pseudolysin Cys^{30}–Cys^{58}. For the C-terminal domain, the template Cys^{270} was aligned with Msp Cys^{510} and Cys^{297} was just three residues away from Cys^{540} located in the flexible terminal loop. There was no previously published evidence of the presence of a structural Ca^{2+} ion in Msp. This was structurally confirmed by our modeling study that shows that three of the five charged residues involved in Ca^{2+} coordination in pseudolysin are replaced by neutral residues in Msp, thus precluding binding of the charged calcium ion.

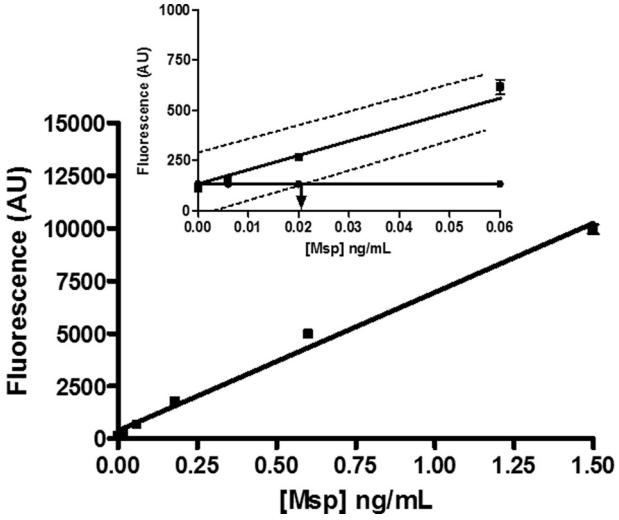

FIGURE 5. **Msp calibration curve using compound 2, Ac-SKG-Pya-(3-NO_2)YGGK-NH_2.** Decreasing concentrations of purified Msp were incubated 60 min at 37 °C using the optimal compound **2** (10 μM) substrate in a final volume of 100 μl in 50 mM HEPES, pH 7. End point fluorescence was measured as described under "Experimental Procedures" (λ_{ex} = 340 nm and λ_{em} = 405 nm). The *inset* shows a zoom on the lower Msp concentrations. Indicated is the calculated limit of detection (0.006 ng/ml) whose intersection with the 90% confidence interval standard curves represents the limit of quantification, thus estimated at 0.02 ng/ml of Msp.

Docking of HPI in the Msp Model

We first verified that Gold was able to accurately position HPI in the active site of pseudolysin in the presence of the Zn^{2+} cation. Considering the similarities between the two sites, the position of HPI in the active site of Msp was very similar, with the dipeptide Phe-Asp-NH_2 interacting with the S_1' and S_2' subsites (data not shown). The benzyl moiety is found inside the S_1' hydrophobic pocket lined by Leu^{369}, Gly^{373}, Val^{374}, Ile^{423}, Gly^{424}, Ile^{427}, and the guanidinium of Arg^{438}. Both the CO and the NH_2 of Asn^{349} are hydrogen-bonded to the HPI amide group. The side chain of Asp of HPI was exposed to the solvent. The carboxylate group of the 3-phenyl propionate moiety complexes the catalytic zinc ion and the phenethyl chain interacts with the putative S_1 subsite (Fig. 8). The S_1 subsite of pseudolysin had been defined by analogy with that of vibriolysin (24) and included Asn^{112}, Tyr^{114}, and Tyr^{155} (Asn^{118}, Phe^{120}, and Tyr^{161} in vibriolysin). The sequence alignment between pseudolysin and Msp allowed the Asn^{349}, Tyr^{351}, and Tyr^{392} residues to be proposed as delineating the S_1 pocket of Msp. This S_1 subsite had a large surface exposed to the solvent.

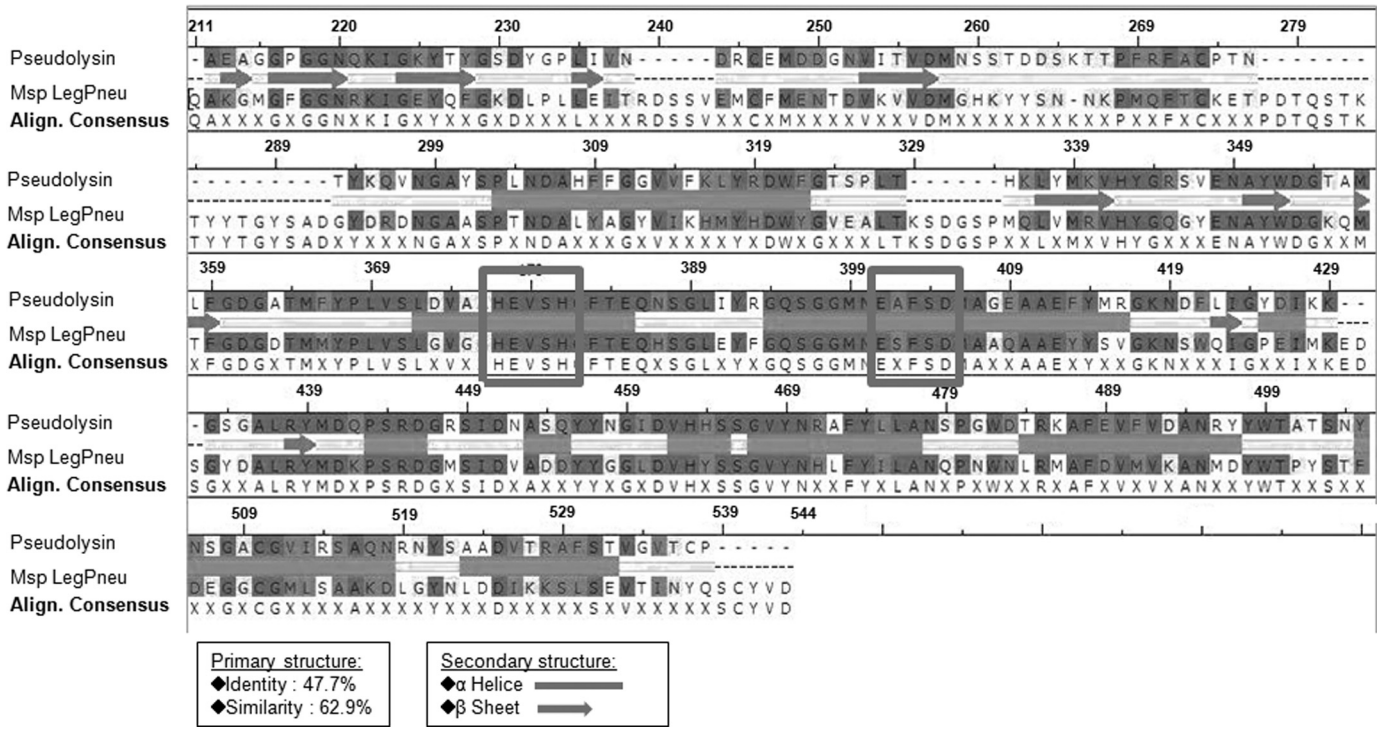

FIGURE 6. **Primary sequence comparison of Msp and pseudolysin.** Sequence alignment of mature pseudolysin (*top*) and of *L. pneumophila* Msp (*Msp LegPneu*, *bottom*) was obtained using ClustalW. The sequence numbers correspond to *L. pneumophila* Msp. The resulting consensus sequence shows identical residues shared by both proteases with *X* representing nonconserved residues. The Msp shares 62.9% sequence similarity and 47.7% identity with pseudolysin. The zinc-binding consensus HE*XX*H and E*XXX*D motifs are *boxed*.

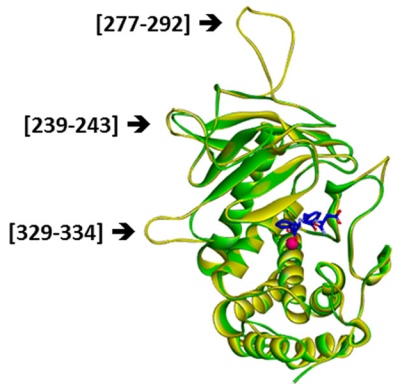

FIGURE 7. **Superimposing of Msp and pseudolysin structures.** The Msp model (*yellow*) obtained with MODELLER is shown superimposed on the pseudolysin crystal structure from the Protein Data Base (*green*). The active site zinc cation is in *pink* and the co-crystallized HPI inhibitor in pseudolysin in *blue*. The *arrows* point to the three Msp loops.

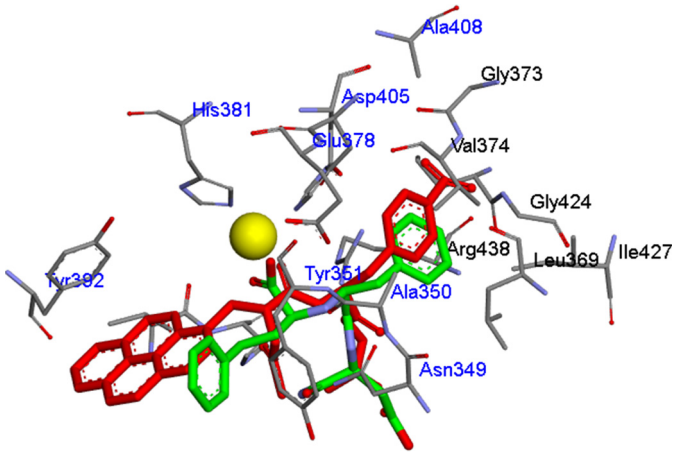

FIGURE 8. **Docking of Ac-Pya-Nop-Gly and HPI in the active site of Msp model structure.** The HPI inhibitor co-crystallized in pseudolysin (*green*) and Ac-Pya-Nop-Gly (*red*) were docked in the Msp model using Gold software. In *black* are the amino acids lining the S_1' subsite and in *blue* those located at a maximum of 6 Å from the active site zinc atom (*yellow*).

Docking of Compound 2 in Model Structure of Msp

In our model, Nop occupied the S_1' subsite; the NH amide was hydrogen bonded to the $CONH_2$ of Asn^{349}, and the NO_2 group interacted with Gly^{424}. The P_2' Gly residue was located in the S_2' subsite, and the CO of Pya bound the zinc ion, and consequently the lateral chain of Pya occupied the S_1 subsite (Fig. 9). In this model, the residues interacting with the S_1 and S_2' subsites lie outside the protein backbone, solvent-exposed. However, the position of the Pya side chain was compatible with an interaction with the aromatic Tyr^{392}, a residue lining the S_1 subsite at the surface of the protein.

This preliminary model was further validated by the results of the enzymatic experiments. These results demonstrated that Msp cleaved compound **1** between Pya and Nop.

Compound **2** was docked using the constraints extracted from the Ac-Pya-Nop-Gly/Msp complex. For the 3-NO_2-Tyr moiety, the amide NH created a hydrogen bond with the CO of Ala^{350} and the carbonyl side chain of Asn^{349}. The phenolic group was positioned inside the hydrophobic S_1' subsite, and the NO_2 group interacted with the guanidinium of Arg^{438}, which lines this subsite.

The Ac-SKG and GGK-NH_2 tripeptides, surrounding the Pya-(3-NO_2)-Tyr core, are highly flexible and allow for various interactions with residues at the surface of the protein. For the C-terminal GGK tripeptide, the two glycine residues induced a β-turn structure thus allowing an ionic interaction between the ϵNH_2 of the C-terminal Lys and the side chain carboxylate of Asp^{446}. The N-terminal Ac-SKG also lies outside the protein, but it could form hydrogen bonds between the Ser-OH and α-COO^- of Asp^{353}, between the Lys α-CO and OH of Tyr-382 OH, and between the Lys ϵNH_2 and COO^- of Glu^{385} (Fig. 9).

DISCUSSION

L. pneumophila is a virulent human pathogen, and the rapid detection of its presence in different water circuits is an important stake for public health. To date, there is no rapid method available for the detection and/or quantification of this pathogen.

Consequently, we studied the possibility of developing a rapid detection of water contamination by monitoring a soluble marker of the bacteria. With this aim, we selected the *L. pneumophila* endopeptidase Msp, which is released in the culture medium of these bacteria in large quantities and had previously been suggested as specific for this family (5), as a possible marker. To develop a rapid detection method for the Msp protease, we searched for a highly fluorescent and selective substrate.

Biochemical and structural data concurred in the design of an optimized substrate. Screening the Fluofast library, the main physicochemical characteristics of this protease were determined. Indeed, the screening unraveled the first critical parameter, *i.e.* the ability of Msp to cleave the Ac-SKG-Pya ↓ Nop-GGK-NH_2 peptide between the two bulky non-natural amino acids Pya and Nop, all the other sub-libraries, possessing one or more residues between this fluorophore/quencher pair, being significantly less or not at all cleaved. From this initial result, it could be concluded that Pya in the S_1 subsite and Nop in the S_1' subsite were ideally positioned in the Msp active site. Based on the structural model of Msp, we then demonstrated that Nop could be favorably replaced by the 3-NO_2-Tyr quencher moiety, the latter showing optimized stabilizing interactions into the S_1' subsite. Interestingly, cleavage of this substrate by Msp was found to yield a signal increased by 2-fold, with the same fluorescent metabolite (Ac-SKG-Pya), indicating the different inter-molecular quenching efficiencies of the nonfluorescent metabolites (Fig. 3).

The N- and C-terminal sequences SKG and GGK, introduced in the library general formula for both increased solubility and ionization in LC/MS detection, were fortuitously found to play important functions in protease binding. Indeed the position and nature of these residues in the substrates were found to modulate their binding to and cleavage by Msp. This is, for example, the case of the positively charged C-terminal Lys, Orn, Arg, or hArg residues of compounds **2, 5, 6, 8,** and **9**. Indeed, the ensemble of the results obtained show that introduction of homo-arginine in the P_2' position allowed for the highest affinity (K_m = 1.1 ± 0.2 μM). Likewise the absence of the N-terminal lysine in compound **13** completely abolished the cleavage of the peptide and replacement of Ser by hSer

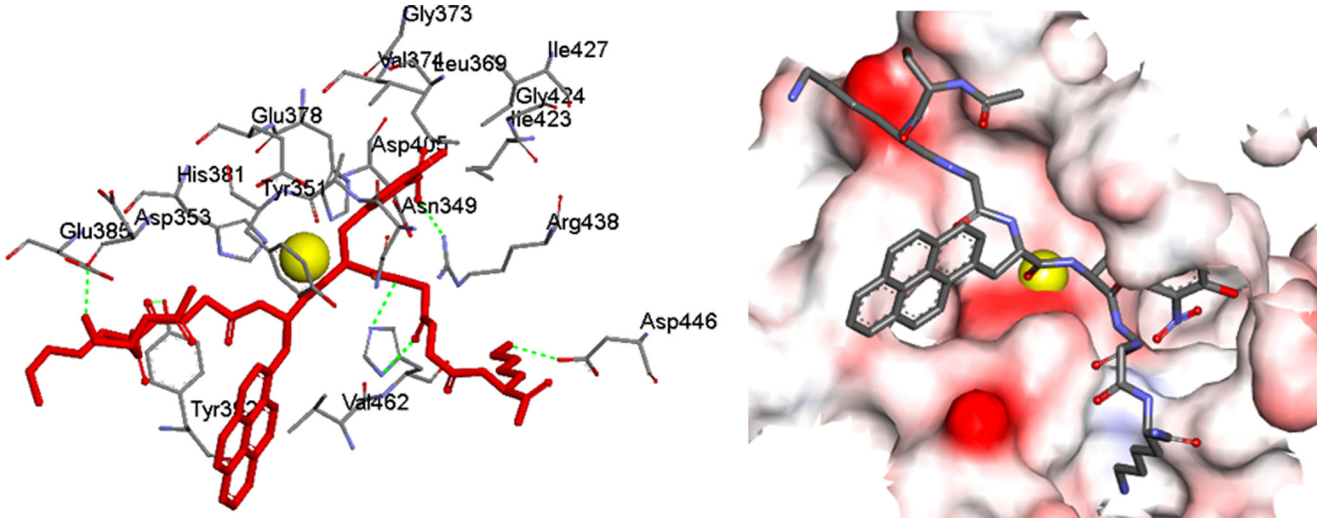

FIGURE 9. **Docking of compound 2 (Ac-SKG-Pya-(3-NO_2)YGGK-NH_2) in the Msp active site.** *Left,* compound **2** (*red*) was docked in the Msp active site model using Gold. The zinc ion is represented in *yellow*. The CO of the N-terminal lysine residue is shown to form a hydrogen bond with Tyr^{392}; the serine forms a hydrogen bond with Asp^{353}, and the (3-NO_2)Tyr quencher moiety lies within the S_1' subsite, interacting with Ala^{350}, Asn^{349}, and Arg^{438}. The C-terminal Lysine residue is hydrogen bonded to Asp^{446} of the Msp protein backbone. *Right,* representation of the Msp protein structure with compound **2** docked within its active site. Although the (3-NO_2)Tyr quencher moiety is fitted inside the S_1' subsite, the Pya fluorophore is shown to lie outside the protein structure, readily soluble upon cleavage of the Pya ↓ (3-NO_2)Tyr scissile bond.

(compounds **6** and **16**) or its suppression (compounds **17** and **18**) modified the properties of the substrate.

The high sequence homology between Msp and pseudolysin allowed the construction of a three-dimensional model of Msp. This strategy was previously used for the modeling of vibriolysin, another member of the M4 peptidase family (24), which possesses 62% sequence identity with pseudolysin and 49.2% with Msp. The overall tertiary structure proposed for mature Msp strongly resembles that of mature pseudolysin, *e.g.* the catalytic sites (zinc ligands and stabilization of the transition state by a Tyr^{392} and a His^{463}) are identically organized, and the S_1 and S_1' sub-sites are both hydrophobic (Figs. 6 and 7).

However, despite these structural homologies, both proteases present very distinct cleavage profiles of the Fluofast library peptides. Indeed, although Msp almost exclusively cleaved Ac-SKG-Pya ↓ Nop-GGK-NH_2 (compound **1**), pseudolysin was unable to cleave the latter but cleaved all the sublibraries with $n = 1$, 2, or 3. To better understand this opposite behavior, their structures were examined in more detail.

Two main differences may be underlined. First, as compared with pseudolysin sequence, three insertion domains were identified in the N-terminal part of Msp (Figs. 6 and 7). These sequences left as nonoptimized structures are located at the protein surface and presumably have no function in Msp enzymatic activity. The herein proposed alignment differs from the one previously published by Black *et al.* (8) who suggested the second loop (here amino acids 277–292) to be constituted by amino acids 259–273. This proposition includes a single cysteine residue in the Msp 259–273 loop, thus abolishing the formation of a disulfide bond found in the pseudolysin structure and otherwise present in Msp. As it most certainly participates in the three-dimensional structural stabilization of the protease, it appears that our model may better reflect the physiological conformation of Msp.

Second, Msp does not possess a structural Ca^{2+} site. In pseudolysin, the Ca^{2+} ion is chelated by residues Asp^{136}, Glu^{172}, Glu^{175}, Asp^{183}, and Leu^{185} found in the helical domain containing the Zn^{2+} ligands (His^{140}, His^{144}, and Glu^{164}) and the residues involved in the catalytic process (such as Glu^{141} and Tyr^{155}). The presence of this cation in close proximity to the active site imposes a constraint that may have incidences particularly on the structure of the S_1' subsite of pseudolysin as compared with Msp.

This latter aspect is brought to light by the fact that the S_1' subsite of pseudolysin allowed the binding of aromatic residues, such as Phe or Tyr side chains, but was unable to accept the NO_2-Phe side chain of compound **1** nor the 3-NO_2-Tyr of compounds **2** and **4**-**19.** An explanation to this result is provided by the structural analysis of Msp. Indeed, modeling of Ac-Pya-Nop-Gly in pseudolysin, with Nop in the S_1' subsite, revealed a steric hindrance and most probably Van der Waals constraints between the nitro group of Nop and the side chain of Asp-136 found lining the bottom of the S_1' pocket of pseudolysin. In these conditions, substrates adopt another position in the active site of pseudolysin, as shown by the cleavage of the $n = 1$ sub-library (Ac-SKG-Pya ↓ *X*-Nop-GGK-NH_2) between Pya and *X*, where *X* is aromatic (Phe or Trp) or hydrophobic (Leu or Ile), and Nop was displaced toward the S_2' subsite of the protease.

Interestingly, Asp^{136}, lining the bottom of the S_1' subsite of pseudolysin, is replaced by a glycine in Msp (Gly^{373}), providing a more flexible structure for the S_1' pocket of Msp, which can easily be adapted to fit Nop or 3-nitro-Tyr residues but not the more bulky 3,5-dinitro-Tyr. It is also interesting to observe that the N- and C-terminal surrounding sequences SKG and GGK, which were introduced in the general formula of the library for both increased solubility and ionization in LC/MS detection, could play an important role in the recognition of the protease. Indeed we observed that the position and nature of these residues in the substrates modulate their recognition by Msp.

In conclusion, screening of the Fluofast library allowed the identification of a specific lead substrate for Msp (compound

1). Rational modifications of this lead substrate permitted the identification of a highly sensitive and selective substrate of Msp. The fluorescent assay of the Msp activity using this novel substrate (Ac-SKG-Pya-(3-NO_2)YGGK-NH_2) (compound **2**) allows the detection of 6 pg/ml pure Msp in 60 min. To transform this *in vitro* assay, with purified Msp, into a useful detection test for contaminated water samples from various hot water or cooling tower systems, many parameters now have to be explored, such as the correlation between the measured enzymatic activity and the concentration of bacteria, the stability of the protease in different conditions, and the selectivity toward other pathogens. These experiments have been performed and have led to the design of a "fluorimetric kit" whose sensitivity provides a novel, simple, and rapid field detection system for *Legionella* risk assessment. These results will be published in specialized publications.

Acknowledgments—We thank Nathalie Jullian for help with the modeling and docking processes. We are also grateful for the technical assistance of Prof. Bernard P. Roques (University René Descartes, Paris) and for the critical review of the manuscript. We also thank Suez Environnement for its support and participation in the project.

REFERENCES

1. Dowling, J. N., Saha, A. K., and Glew, R. H. (1992) Virulence factors of the family Legionellaceae. *Microbiol. Rev.* **56,** 32–60
2. Baine, W. B., Rasheed, J. K., Mackel, D. C., Bopp, C. A., Wells, J. G., and Kaufmann, A. F. (1979) Exotoxin activity associated with the Legionnaires disease bacterium. *J. Clin. Microbiol.* **9,** 453–456
3. Müller, H. E. (1980) Proteolytic action of *L. pneumophila* on human serum proteins. *Infect. Immun.* **27,** 51–53
4. Thompson, M. R., Miller, R. D., and Iglewski, B. H. (1981) *In vitro* production of an extracellular protease by *Legionella pneumophila*. *Infect. Immun.* **34,** 299–302
5. Dreyfus, L. A., and Iglewski, B. H. (1986) Purification and characterization of an extracellular protease of *Legionella pneumophila*. *Infect. Immun.* **51,** 736–743
6. Keen, M. G., and Hoffman, P. S. (1989) Characterization of a *Legionella pneumophila* extracellular protease exhibiting hemolytic and cytotoxic activities. *Infect. Immun.* **57,** 732–738
7. Quinn, F. D., and Tompkins, L. S. (1989) Analysis of a cloned sequence of *Legionella pneumophila* encoding a 38-kDa metalloprotease possessing hemolytic and cytotoxic activities. *Mol. Microbiol.* **3,** 797–805
8. Black, W. J., Quinn, F. D., and Tompkins, L. S. (1990) Legionella pneumophila zinc metalloprotease is structurally and functionally homologous to Pseudomonas aeruginosa elastase. *J. Bacteriol.* **172,** 2608–2613
9. Moffat, J. F., Black, W. J., and Tompkins, L. S. (1994) Further molecular characterization of the cloned *Legionella pneumophila* zinc metalloprotease. *Infect. Immun.* **62,** 751–753
10. Berdal, B. P., Olsvik, O., Myhre, S., and Omland, T. (1982) Demonstration of extracellular chymotrypsin-like activity from various *Legionella* species. *J. Clin. Microbiol.* **16,** 452–457
11. Berdal, B. P., Hushovd, O., Olsvik, O., Odegard, O. R., and Bergan, A. T. (1982) Demonstration of extracellular proteolytic enzymes from *Legionella* species strains by using synthetic chromogenic peptide substrates. *APMIS Sect. B* **90,** 119–123
12. McIntyre, M., Quinn, F. D., Fields, P. I., and Berdal, B. P. (1991) Rapid identification of *Legionella pneumophila* zinc metalloprotease using chromogenic detection. *APMIS* **99,** 316–320
13. Thayer, M. M., Flaherty, K. M., and McKay, D. B. (1991) Three-dimensional structure of the elastase of *Pseudomonas aeruginosa* at 1.5-Å resolution. *J. Biol. Chem.* **266,** 2864–2871
14. Roques, B. P., Luciani, N., Fournié-Zaluski, M. C., and de Rocquigny, H. (January 30, 2007) U. S. Patent 7,169,574
15. Poras, H., Ouimet, T., Orng, S. V., Dangé, E., Fournié-Zaluski, M. C., and Roques, B. P. (2011) Pluripotentialities of a quenched fluorescent peptide substrate library. Enzymatic detection, characterization, and isoenzymes differentiation. *Anal. Biochem.* **419,** 95–105
16. Soleilhac, J. M., Cornille, F., Martin, L., Lenoir, C., Fournié-Zaluski, M. C., and Roques, B. P. (1996) A sensitive and rapid fluorescence-based assay for determination of tetanus toxin peptidase activity. *Anal. Biochem.* **241,** 120–127
17. Ristroph, J. D., Hedlund K. W., and Allen, R. G. (1980) Liquid medium for growth of *Legionella pneumophila*. *J. Clin. Microbiol.* **11,** 19–21
18. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Clustal W. Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties, and weight matrix choice. *Nucleic Acids Res.* **22,** 4673–4680
19. Sali, A., and Blundell, T. L. (1993) Comparative protein modeling by satisfaction of spatial restraints. *J. Mol. Biol.* **234,** 779–815
20. Ghose, A. K., and Viswanadhan, V. N. (2001) *Combinatorial Library Design and Evaluation. Principles, Software, Tools, and Applications in Drug Discovery*, CRC Press, Inc., New York
21. Jones, G., Willett, P., Glen, R. C., Leach, A. R., and Taylor, R. (1997) Development and validation of a genetic algorithm for flexible docking. *J. Mol. Biol.* **267,** 727–748
22. Meldal, M., Svendsen, I., Breddam, K., and Auzanneau, F. I. (1994) Portion-mixing peptide libraries of quenched fluorogenic substrates for complete subsite mapping of endoprotease specificity. *Proc. Natl. Acad. Sci. U.S.A.* **91,** 3314–3318
23. Schechter, I., and Berger, A. (1967) On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* **27,** 157–162
24. Lutfullah, G., Amin, F., Khan, Z., Azhar, N., Azim, M. K., Noor, S., and Shoukat, K. (2008) Homology modeling of hemagglutinin/protease (HA/P (vibriolysin)) from *Vibrio cholerae*. Sequence comparison, residue interactions, and molecular mechanism. *Protein J.* **27,** 105–114