Characterisation of an extracellular serine protease gene (nasp gene) from *Dermatophilus congolensis*

Alfredo Garcia-Sancheza, Rosario Cerrato, Jose Larrasa, Nicholas C. Ambrose, Alberto Parra, Juan M. Alonso, Miguel Hermoso-de-Mendoza, Joaquin M. Rey, Madisa O. Mine, Patrick R. Carnegie, Trevor M. Ellis, Anne M. Masters, Alan D. Pemberton, Javier Hermoso-de-Mendoza

**Abstract**

A partial amino acid sequence of a serine protease from *Dermatophilus congolensis* allowed the design of oligonucleotide primers that were complemented with additional ones from previously published partial sequences of the gene encoding the enzyme. The polymerase chain reaction (PCR), using combinations of specific and degenerate oligonucleotide primers, allowed the amplification of a 1738-bp internal fragment of the gene, which was finally characterised by inverse PCR as the first full-length sequenced serine protease gene (nasp) from *Dermatophilus congolensis*. The deduced amino acid sequence of this enzyme, probably involved in the pathogenesis of dermatophilosis, links it to the subtilisin family of proteases.

Keywords: Dermatophilosis; nasp gene; Serine protease; Pathogenesis

1. Introduction

*Dermatophilus congolensis* is the aetiological agent of dermatophilosis, a widely distributed animal disease, occasionally zoonotic, characterised by exudative and proliferative dermatitis. The disease has greater impact when associated with a wet environment and tick infestation [1].

The majority of *D. congolensis* virulence factors have not been defined, however, *D. congolensis* is haemolytic [2] and produces phospholipases [3], ceramidases (Garcia-Sanchez et al., in press) and proteolytic enzymes [4–6]. Dermatophilosis lesions are characterised by inflammation, production of serous exudate and scab formation. *D. congolensis* is normally confined to the epidermis, being unable to reach deeper layers. The cell envelope of epidermal corneocytes is composed of proteins overlain by lipids and the intercellular spaces are lipid-rich [7]. Ceramidases and proteases may be used by *D. congolensis* to penetrate this barrier. Alternatively *D. congolensis* proteases may initiate or inactivate host inflammatory protease cascades, hydrolyse immune effector and regulatory proteins [8] or activate or inactivate cytokines [9].

The lack of detailed information on virulence factors of *D. congolensis* has hindered major progress in understand-
ing the dermatophilosis pathogenesis and in the immunological control of the disease [10]. The use of degenerate primers and heat-soaked polymerase chain reaction (PCR) allowed characterisation of partial nucleotide sequences of serine protease genes from two sheep isolates of *D. congolensis* [11]. In a later study of *D. congolensis* extracellular products, the production of extracellular alkaline serine proteases by this actinomycete was demonstrated [12].

Serine proteases are proteolytic enzymes that depend on a serine residue for catalytic activity. These enzymes have been implicated in the pathogenesis of a number of diseases caused by bacteria and parasites which have to penetrate skin [13–17].

The objective of this study was to characterise the complete nucleotide sequence of an alkaline serine protease of *D. congolensis*, using a partial N-terminal nucleotide sequence of the enzyme and a previously published partial sequence of a serine protease gene [11] as starting points. The present paper is the result of an international coordinated effort of three research teams. The chronic lack of funds for research on this disease is delaying any progress in control, specially urgent in tropical developing countries where the disease is an important livestock problem, frequently aggravating poverty and famine.

### 2. Materials and methods

#### 2.1. Partial sequencing of the serine protease

* * * * *  

*D. congolensis* (ATCC 14637) was grown in 20 ml of 2% w/v brain heart infusion, 1.7% w/v neutralised soya peptone (Oxoid) in glass McCartney bottles, at 37°C in air for 48 h. The culture was then centrifuged and concentrated following a previously published protocol [6]. Separation of extracellular products, characterisation of serine proteases by specific inhibitors (PMSF, DCI and Pefabloc SC®) on substrate sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) and selection of a 24-kDa enzymically active band were performed following previously published protocols [12]. The 24-kDa protease was purified by high performance liquid chromatography, anion exchange and size exclusion chromatography, Tris-Tricine SDS–PAGE, and blotting with CAPS buffer onto polyvinylidene difluoride membrane, and sequenced by Edman degradation.

#### 2.2. Extraction of genomic DNA

* * * * *  

*D. congolensis* (Australian ovine strain MG) was grown as above.

DNA was isolated using a modified version of the protocol described by Wilson [18] with an additional step of 5 h incubation at 37°C with lysozyme (0.2 mg ml⁻¹) in Tris-EDTA buffer, prior to a SDS and proteinase K treatment step, followed by standard phenolic extraction [19].

#### 2.3. Design of oligonucleotide primers and PCR conditions

Our first amino acid partial sequence, (A/E)SP-KAVREKPLXTIGKVFFTRNxED, allowed us to choose a shorter internal peptide (the underlined KAVREKPL) to design the upstream sense degenerate primer SERINPRO1 (5’-AARGCNTGNGARAAACCCNCTN-3’).

The downstream antisense primer was designed from the highly conserved area around the serine protease active site present in bacterial serine proteases (Fig. 1) as previously described [11]. This primer was called SERINPRO2 (3’-CCNTGWNWTACGNDCNGGN-5’).

The PCR with degenerate primers was carried out under various annealing temperatures (40–60°C) using genomic DNA of *D. congolensis* MG strain as template. However, no specific amplified products were obtained, and therefore each degenerate primer was then tested in combination with a specific primer designed on the basis of the partial nucleotide sequence obtained by Mine and Carnege [11] (GenBank AF010468). The primers were called AUSTRALIA 1 (5’-GATGGAATGTCAGGGAGCAC-G-3’) and AUSTRALIA 2 (5’-GTCTTCCGGGCTCCAGAAACAT-3’). The use of these specific primers was expected to achieve higher specificity, as confirmed below.

PCR amplification was performed in 100 µl of reaction mixture which contained 100 ng of genomic DNA, 1.5 mM MgCl₂, 1 µM of each primer, 250 µM (each) deoxynucleoside triphosphate (dNTP), 1.5 U Taq DNA polymerase (BioTaq®), 10 mM Tris- HCl pH 8, 50 mM KCl and 0.1% Triton X-100. PCR conditions were programmed for denaturation at 94°C for 5 min and then for 35 cycles of 1 min at 94°C, 1 min at 55°C, and 1.5 min at 72°C. The high-temperature annealing favoured the specific primers for priming. The fragments obtained were sequenced as explained below, and the resulting nucleotide sequence was used to design the primers for the inverse PCR.

#### 2.4. Inverse PCR

Inverse PCR was performed by treating 2 µg of *D. congolensis* genomic DNA with different restriction enzymes, namely *Pst*I, *Sac*I, *EcoRI*, *Bgl*I and *Sma*I. The restriction endonucleases were then thermally inactivated (80°C, 20 min), and the cut DNA incubated with T4 DNA ligase (Amersham Biosciences) overnight at 16°C. The resulting self-ligated DNA was precipitated with ethanol and resuspended in sterile distilled water to give a final concentration of 10 ng µl⁻¹. This was used as the template in PCR with primers SERINVPOR1 (5’-TCGGGAGACCCTCTCTGTGGGT-3’) and SERINVPOR2 (5’-TGCAAGCCACCGCAAAACGACC-3’) oriented in inverted tail-to-tail direction and designed according to the known nucleotide sequence.

PCR reactions (100 µl) contained 5 or 10 µl self-ligated chromosomal DNA, 0.5 mM of each primer, 0.2 mM
concentrations of each deoxynucleotide triphosphate, 1.5 IU Taq DNA polymerase (BioTaq®, Bioline), 1.5 mM MgCl₂, 10 mM Tris–HCl pH 8, 50 mM KCl and 0.1% Triton X-100. A thermocycler was programmed as follows: 5 min warm-up/hold at 94°C, followed by 30 cycles, each comprising 1 min at 94°C, 1 min at annealing temperature (55°C) and 2 min at 72°C, followed by a 15-min hold at 72°C. PCR products were cloned into vector pCR2.1-TOPO (Invitrogen) and sequenced with vector-specific and custom-designed primers.

2.5. Cloning and sequencing

The PCR products were purified from the 1% agarose gel using QIAquick Gel Extraction kit (Qiagen) and cloned using a TOPO cloning kit (Invitrogen). Randomly selected white colonies were purified using UltraClean Mini Plasmid Prep Kit (MoBio) and the plasmid DNA was sequenced using an ABI 373 autosequencer (Perkin-Elmer) with four fluorescent dyes. Computer-assisted analysis of the sequence data was performed using ABI Prism sequencing analysis version 3.3 (Perkin-Elmer Software).

2.6. Computer sequence analysis

Nucleotide sequence data were compiled using the DNA Strider 1.0 software [20]. Database searches were performed using BlastX (http://www.ncbi.nlm.nih.gov/BLAST [21]). Open reading frames (ORFs) were identified using ORF Finder (http://www.ncbi.nlm.nih.gov/gorf/gorf.html). Signal sequence prediction was performed using SignalP (http://www.cbs.dtu.dk/services/SignalP) [22]. Multiple sequence alignments were performed using PILE-UP (GCG software, version 9.0, [23]). Protein masses and isoelectric points were determined online, using ProtParam tools (http://us.expasy.org/tools/protparam.html). Proteins were examined for conserved motifs using Pfam (http://pfam.wustl.edu/hmmsearch.shtml).

3. Results

The amino acid sequence of the purified serine protease was: N-terminal-(A/E)SPKAVREKPlxTIGKVFFTRN-xED.

Using the degenerate primer SERINP01 (forward) in combination with the specific primer AUSTRALIA 2 (reverse), an approximately 1500-bp PCR product was obtained (Fig. 2). In the same way, the specific primer AUSTRALIA 1 (forward) and the degenerate primer SERINP02 (reverse) were successfully used in the amplification of an approximately 650-bp PCR fragment (Fig. 2). The two PCR products were then cloned into the vector pCR2.1-TOPO (Invitrogen) and sequenced.

The resulting sequences of the two fragments overlapped, consisting of a single sequence of 1738 bp long, which was subsequently used to design new primers for use in inverse PCR (SERINP01 and SERINP02). The sequence analysis showed the PCR product to contain an incomplete ORF and revealed extensive homology (50% identity, 64% similarity) with a putative serine protease gene of *Streptomyces coelicolor* A3 in the EMBL/GenBank Data Library (AL138662).

Finally, successive inverse PCR experiments allowed amplification of the full-length *nasp* gene. Self-ligated circular DNA derived from SacI-digested MG chromosomal DNA was used as a template to finally amplify an approximately 2.5-kb fragment, which was also cloned and sequenced.

The sequences were then aligned and assembled obtaining a final sequence of 3709 bp long.

The nucleotide sequence of 3709 bp comprised an ORF, the *nasp* gene, the most likely start codon (ATG) being located at nucleotide 207, and the stop codon (TAA) located at nucleotide 3481. This *nasp* gene would encode a...
serine protease with 1094 amino acid residues and \( M_r \) of 116,572.

Computer analysis of the deduced protein product of \( nasp \) with the BLAST program identified similarities with the putative serine protease of \( Streptomyces avermitilis \) MA-4680 (NP826577.1) and the putative serine protease of \( S. coelicolor \) A3 (NP626873.1). They respectively have 43% and 42% sequence identity over the full length. The \( nasp \) gene sequence was deposited in GenBank, under accession number AJ459111.

4. Discussion

Degenerate oligo-primers, partial amino acid sequence and inverse PCR have been successfully used to clone the \( nasp \) gene from \( D. congolensis \). This gene would encode a secreted protease that is related to the subtilisin family of proteases. Some features of the deduced amino acid sequence suggest a pathogenic role of the enzyme, namely a putative signal peptide sequence of 30 residues at the N-terminus indicating that it could be secreted, with a theoretical isoelectric point of 7.17. Furthermore, a predicted bacterial pre-peptidase C-terminal domain (Pfam family PF04151) of about 80 amino acids was present in the C-terminal region. So probably there is further processing at the C-terminus to remove this domain.

Typical motifs are observed around the catalytic residues Asp, His and Ser, and around the oxyanion hole Asn (Siezen, personal communication). The high homology with members of the \( Streptomycyes \) genus is not surprising since \( D. congolensis \) is taxonomically related to them in the Actinomycetales. According to Siezen et al. [24], as is true of nearly all enzymes in the subtilase family, the serine protease is synthesised as a pre-pro-enzyme, and is subsequently translocated over the membrane where it is activated by cleavage of the pro-segment.

There is evidence of important extracellular protease activity in \( D. congolensis \) infection [4–6]. The invasive phase of its life cycle takes advantage of such activity, in which serine proteases are involved [12], as well as other favourable enzyme activities, e.g. ceramidases (Garcia-Sanchez et al., in press), to break the main protective barriers like keratin and cutaneous lipids, allowing penetration and survival in deeper epidermal layers, and the completion of its life cycle [1].

This study will lead to a better understanding of the host–\( Dermatophilus \) relationship and identifies a virulence factor that could be considered a candidate antigen for vaccines and a diagnostic tool for dermatophiosis. Its use as an immunising antigen might prevent the ability of \( D. congolensis \) to infect the skin of animals, as a complementary measure to tick control and appropriate management in countries affected by dermatophiosis. Further studies must be carried out for recombinant expression of this and other enzymes, in order to assess their usefulness for development of therapeutic or immunological interventions to control dermatophiosis.

Acknowledgements

This work was supported by a Grant from the European Union programme INCO-DC ‘Genetic and immunological control on dermatophiosis’, Contract IC18-CT98-0334 (DGXII-SNRD). The authors are grateful to Julia M. Carson for her collaboration in Australian strain typing. Special thanks to Dr Alberto Quesada, Department of Biochemistry, University of Extremadura, for his valuable suggestions for the characterisation of the gene, and to Dr Roland Siezen, Centre for Molecular Biology and Molecular Informatics, University of Nijmegen, The Netherlands, for his generous contribution in the analysis of the \( nasp \) gene sequence, suggestions and predictions on the 3D structure of the \( nasp \) serine protease.

References


\( \text{tophillus congolensis} \). \text{Zentbl. Vet.med.} \text{B} 39, 139–143.


\( tophillus congolensis \) measured by API ZYM. Vet. Microbiol. 37, 175–179.


