Biochemistry and Biophysics Reports 5 (2016) 448–452

Contents lists available at ScienceDirect

Biochemistry and Biophysics Reports

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

Expression and production of soluble *Mycobacterium tuberculosis* H37Rv mycosin-3

Zhuo Fang [a,*], Wolf-Dieter Schubert [b], Nicolaas Claudius Gey van Pittius [a]

[a] DST/NRF Centre of Excellence in Biomedical Tuberculosis Research, US/MRC Centre for Molecular and Cellular Biology, Division of Molecular Biology and Human Genetics, Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, University of Stellenbosch, Francie van Zijl Drive, Tygerberg 7505, South Africa

[b] Department of Biochemistry, University of Pretoria, cnr Lynnwood Road and Roper Street, Hatfield, Pretoria 0028, South Africa

ARTICLE INFO

Article history:
Received 15 December 2015
Received in revised form
3 February 2016
Accepted 4 February 2016
Available online 6 February 2016

Keywords:
Mycobacterium tuberculosis
ESX
Mycosin-3
Protein production
Cloning

ABSTRACT

Mycobacteria encode five type VII secretion system (T7SS) or ESX for nutrient acquisition and virulence. Mycosins are membrane-anchored components of ESX with serine protease activity but an unidentified substrate range. Establishing the substrate specificity of individual mycosins will help to elucidate individual ESX functions. Mycosin-1 and -3 orthologues from two environmental mycobacterial species, *Mycobacterium smegmatis* and *Mycobacterium thermoresistibile*, have been heterologously produced, but mycosins from *Mycobacterium tuberculosis* (*Mtb*) remain to be studied. Here we describe the successful production of *Mtb* mycosin-3 as a first step in investigating its structure and function.

1. Introduction

Pathogenic Gram negative bacteria use a range of secretion systems to secrete virulence factors or transport the factors into host cells to manipulate the host immune system [1]. Type VII secretion systems (T7SSs) are restricted to mycobacteria and some other high GC Gram-positive bacteria [2,3]. *Mycobacterium tuberculosis* (*Mtb*), the etiological agent of tuberculosis, has five T7SSs, denoted as ESX-1 to -5 presumably evolved by gene duplication [4]. ESX-1 and -5 are critical to virulence in pathogenic mycobacteria [5], and ESX-3 participates in mycobactin-mediated iron acquisition [6,7]. ESX-5 was recently found to additionally function in nutrient acquisition [8]. ESX-1, 3 and 5 are correspondingly essential for *Mtb* growth *in vitro* [9,10]. The roles of ESX-2 and -4 are not yet clear. The close association of ESXs with fundamental biological processes has resulted in much research interest in T7SS.

Details of T7SS secretion have not been fully elucidated including the highly conserved mycosin components. Analysing mycosins may therefore help to unravel their functions. Mycosin-5 was not co-isolated with the central, double membrane spanning complex consisting of EccB, EccC, EccD and EccE, indicating a weak association *in vivo* [11]. Mycosins share a conserved catalytic triad of aspartate, histidine and serine with subtilisin-like serine proteases [12]. Screening experiments, however, did not identify mycosin substrates [12]. Recently, mycosin-1 ($MycP_1$) was found to cleave EspB twice upon secretion [13] to potentially facilitate its maturation for host target interaction. This is, however, unlikely to be the only mycosin substrate, as the gene *espB* is unique to ESX-1. ESX-1 substrate secretion is dependent on mycosin-1 but removing its enzymatic activity unexpectedly increases secretion [13]. Mycosin-1 may thus ensure *Mtb* persistence by balancing immune detection and virulence [13].

Mycosins have an N-terminal secretion signal followed by a potential “pro-peptide”, a catalytic domain, a proline-rich linker and a hydrophobic transmembrane region (Fig. 1). While removal of the “pro-peptide” was originally proposed to be required for enzymatic activation [12], this was found not to affect its protease activity [14–16]. In addition, crystal structures of mycosin-1 from *M. smegmatis* and *M. thermoresistibile* and mycosin-3 ($MycP_3$) from *M. smegmatis* suggest that the “pro-peptide” wraps around the catalytic domain possibly to stabilize it. The “pro-peptide” has hence been renamed the “N-terminal extension region” [14–16]. The $MycP_1$ orthologue of *M. smegmatis* inefficiently cleaves EspB *in vitro* possibly due to other ESX-1 components being absent [14]. However, mycosin-1 orthologues from *M. smegmatis* and *M. thermorsistibile* are unlikely to be involved in virulence in these saprophytic species despite an amino acid sequence identity of 70% with *Mtb* protein. Production of recombinant mycosin-1 or -3 from *M. tuberculosis* is problematic. Although the role of mycosin-3 remains enigmatic, it is essential to *M. tuberculosis* survival *in vitro*

* Corresponding author.
E-mail addresses: zf@sun.ac.za (Z. Fang), wolf-dieter.schubert@up.ac.za (W.-D. Schubert), ngvp@sun.ac.za (N.C. Gey van Pittius).

http://dx.doi.org/10.1016/j.bbrep.2016.02.005

```
1-   MIRAAFACLAATVVVAGWWTPPAWAIGPPVVDAAAQPPSGDPGPVAPMEQRGACSVSGVI
61-  PGTDPGVPTPSQTMLNLPAAWQFSRGEGQLVAIIDTGVQPGPRLPNVDAGGDFVESTDGL
121- TDCDGHGTLVAGIVAGQPGNDGFSGVAPAARLLSIRAMSTKFSPRTSGGDPQLAQATLDV
181- AVLAGAIVHAADLGAKVINVSTITCLPADRMVDQAALGAAIRYAAVDKDAVIVAAAGNTG
241- ASGSVSASCDSNPLTDLSRPDDPRNWAGVTSVSIPSWWQPYVLSVASLTSAGQPSKFSMP
301- GPWVGIAAPGENIASVSNSGDGALANGLPDAHQKLVALSGTSYAAGYVSGVAALVRSRYP
361- GLNATEVVRRLTATAHRGARESSNIVGAGNLDAVAALTWQLPAEPGGGAAPAKPVADPPV
421- PAPKDTTPRNVAFAGAAALSVLVGLTAATVAIARRRREPTE
```

Fig. 1. The primary structure of *M. tuberculosis* mycosin-3 ($MycP_3$). Single underline: signal peptide; double underline: N-terminal extension; dashed underline: proline-rich linker; wave underline: hydrophobic transmembrane region; white on black: catalytic triad, Asp^{95}-His^{126}-Ser^{342}.

Table 1
Primers used to generate the starting *M. tuberculosis* $mycP_3$ construct (pET-28a construct was not codon-optimized), and eight codon-optimized $mycP_3$ constructs (Constructs A to I), Expression hosts and vectors are as listed.

Construct Name	Encoded Amino Acid Sequence	Expression Host	Expression Vector	Primer Sequences and Their Restriction Sites
Construct A	Ile^{26}-Asn^{430}	*E. coli* BL21 (DE3) pLysS	pET-28a	forward: 5′-CCATGGCGATCGGGCCGCCGG-3′ (*Nco*I) reverse: 5′-CTCGAGGTTGCGCGGTGTGGTG-3′ (*Xho*I)
			pGEX-6P-1	N/A (restricted directly from the synthetic construct)
Construct B	Arg^{51}-Asn^{430}		pET-28a	forward: 5′-CCATGGAACGCGGTGCGTGCAG-3′ (*Nco*I) Construct A reverse primer
			pGEX-6P-1	forward: 5′-GGATCCCGCGGTGCGTGCAG-3′ (*BamH*I) Construct A reverse primer
Construct C	Gly^{52}-Leu^{401}		pGEX-6P-1	forward: 5′-GGATCCGGTGCATGTAGCG-3′ (*BamH*I) reverse: 5′-CTCGAGTCACAGCTGCCAGGTC-3′ (*Xho*I)
Construct D	Ser^{57}-Leu^{401}		pGEX-6P-1	forward: 5′-GGATCCGGTGTTATTCCGG-3′ (*BamH*I) Construct C reverse primer
Construct E	Gly^{62}-Leu^{401}		pGEX-6P-1	forward: 5′-GGATCCGGTACAGATCCGG-3′ (*BamH*I) Construct C reverse primer
Construct F	Val^{67}-Leu^{401}		pGEX-6P-1	forward: 5′-GGATCCGTTCCGACCCCGAG-3′ (*BamH*I) Construct C reverse primer
Construct G	Ser^{71}-Leu^{401}	*E. coli* Arctic Express and BL21 (DE3) pLysS	pGEX-6P-1	forward: 5′-GGATCCCAGACCATGCTG-3′ (*BamH*I) Construct C reverse primer
			pCOLD	forward: 5′-CATATGCAGACCATGCTGAATC-3′ (*Nde*I) Construct C reverse primer
Construct H	Leu^{77}-Leu^{401}	*E. coli* BL21 (DE3) pLysS	pGEX-6P-1	forward: 5′-GGATCCCTGCCAGCAGCATG-3′ (*BamH*I) Construct C reverse primer
Construct I	Ile^{24}-Leu^{401}	*E. coli* Origami II and Rosetta gami II	pGEX-6P-1	forward: 5′-GGATCCATTGGTCCGCCTGTTG-3′ (*BamH*I) Construct C reverse primer

*Underlined sequences are restriction sites as indicated in brackets.

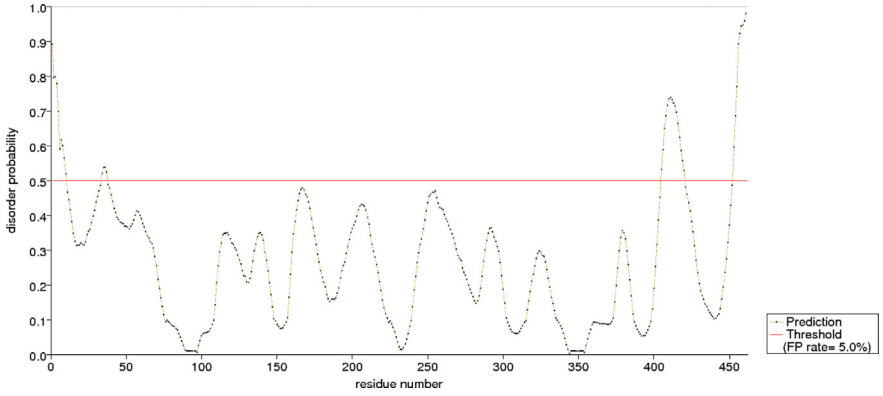

Fig. 2. Disordered region probability plot for *Mtb* H37Rv $MycP_3$ with a prediction false positive rate of 5%, where the disordered region prediction software DrDOS identified two disordered regions, Met^{1}-Pro^{37} and Pro^{405}-E^{461}.

[9,17] making it a potential anti-TB drug target [15,18].

In this study, the gene $mycP_3$ from *M. tuberculosis* H37Rv was cloned and expressed. Extensive effort was made to optimize the construct for soluble mycosin-3 production to increase yield and stability. This report may aid efforts to study mycosin-3 with respect to substrate screening, functional characterization, enzyme kinetics and crystal structure determination.

2. Materials and methods

2.1. Media, plasmids and bacteria strains

Lysogeny broth (LB) was used to culture all *Escherichia coli* strains including XL-1 blue (Promega), BL21 (DE3) pLysS (Promega), Arctic Express (Agilent Technologies), Origami II (Novagen), and Rosetta gami II (Novagen). *E. coli* expression vector pGEX-6P-1