

Supplementary data

Understanding specificity of the mycosin proteases in ESX/type VII secretion by structural and functional analysis

Jonathan W. Wagner¹, Timothy J. Evans¹, Jing Chen¹, Haining Zhu¹, Edith N. G. Houben², Wilbert Bitter² and Konstantin V. Korotkov^{1*}

¹Department of Molecular & Cellular Biochemistry and Center for Structural Biology, University of Kentucky, Lexington, Kentucky 40536, USA

²Department of Medical Microbiology and Infection Control, VU University Medical Center, Amsterdam, The Netherlands

*Correspondence to: Konstantin V. Korotkov, Department of Molecular & Cellular Biochemistry, University of Kentucky, 741 South Limestone, Lexington, KY 40536, USA.
Phone: 859-323-5493. Fax: 859-323-5504. E-mail: kkorotkov@uky.edu

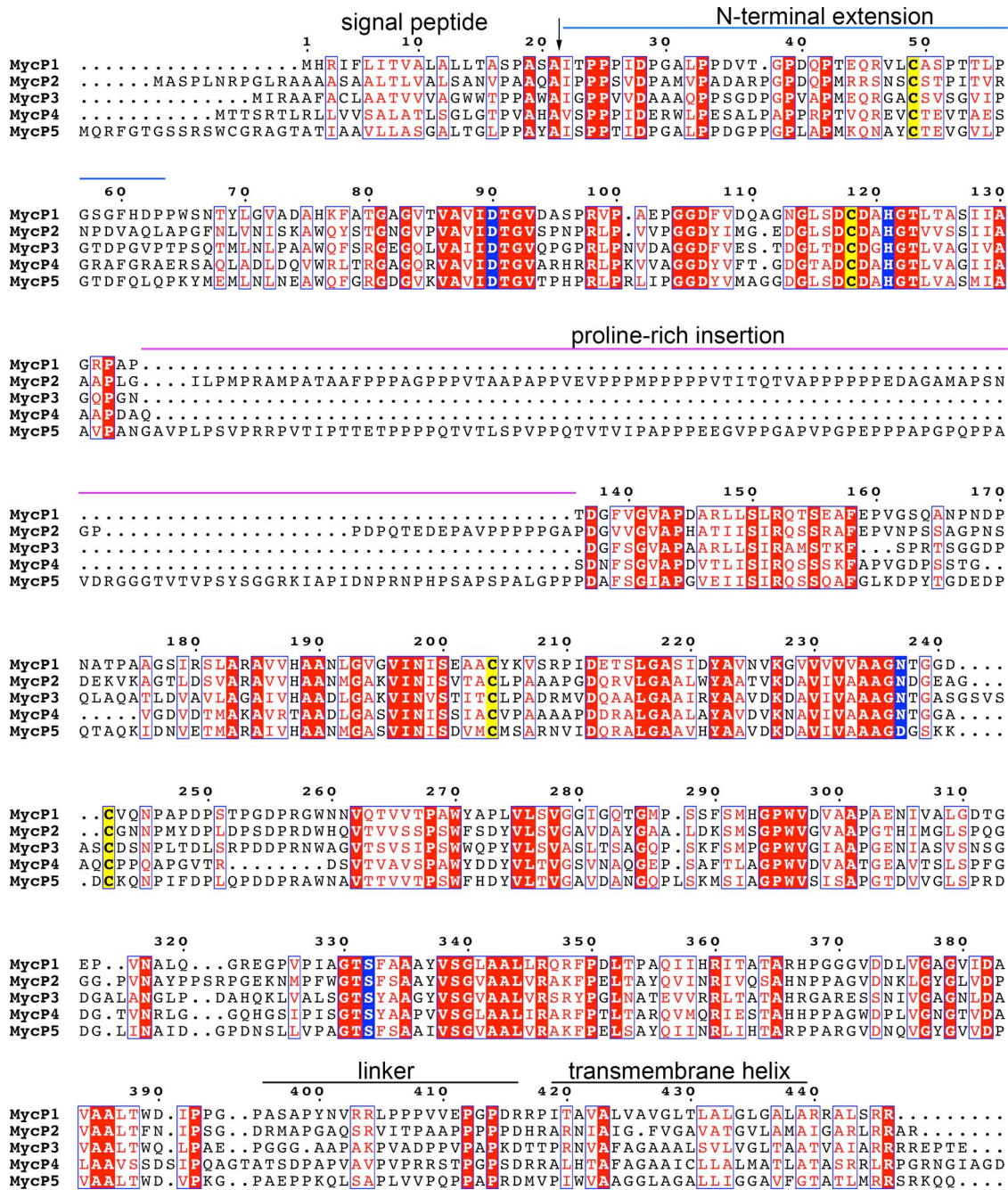


Fig. S1. Sequence alignment of *M. tuberculosis* mycosin proteases.

The N-terminal extension is indicated by blue line. The proline-rich insertion, which is present in MycP₂ and MycP₅, is indicated by magenta line. The conserved Cys residues are highlighted in yellow. The catalytic triad residues, as well as the residues coordinating the oxyanion hole, are highlighted in blue. Note that Asp367 is the residue coordinating the oxyanion hole in MycP₅.

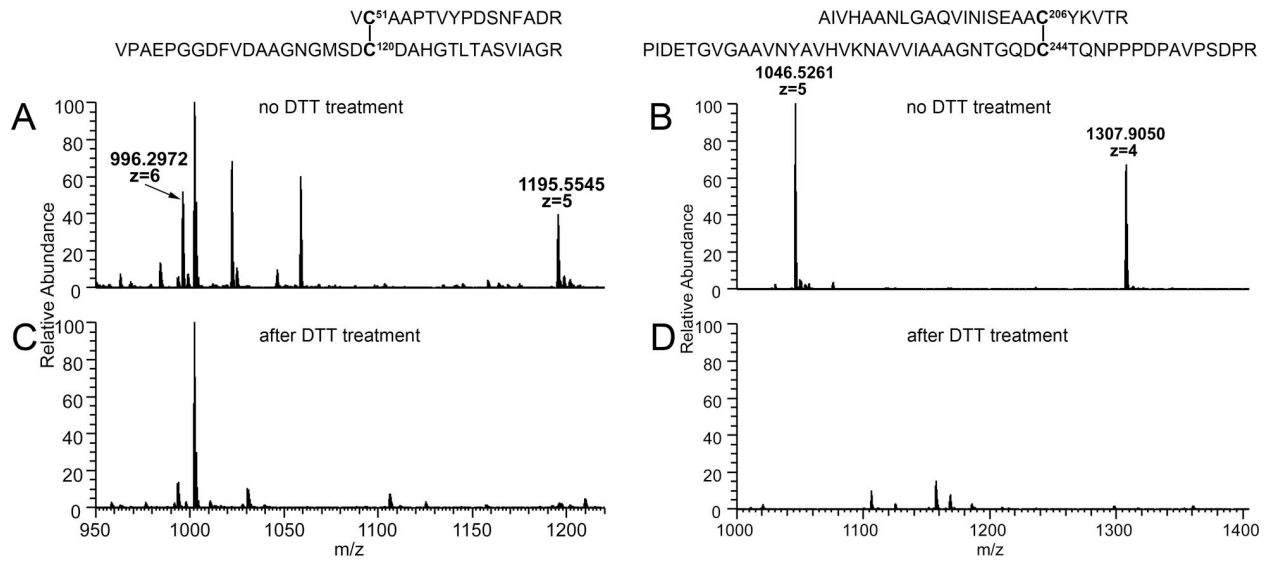


Fig. S2. Mass spectrometry analysis of the disulfide-linked peptides from MycP_{1mth}.

A and B. MycP_{1mth} sample was not pretreated with dithiothreitol (DTT) and iodoacetamide (IAA) prior to trypsin digestion.

C and D. MycP_{1mth} sample was pretreated with DTT and IAA prior to trypsin digestion.

The peaks labeled in (A) are the 6+ and 5+ ions of the disulfide-linked peptides: VC⁵¹AAPTVPDSNFADR–VPAEPGGDFVDAAGNGMSDC¹²⁰DAHGTLTASVIAGR. The peaks

labeled in (B) are the 5+ and 4+ ions of the disulfide-linked peptides:

AIVHAANLGAQVINISEAAC²⁰⁶YKVTR–

PIDETGVGAAVNYAVHVKNNAVIAAAGNTGQDC²⁴⁴TQNPPDPAPVPSDPR. These peaks disappeared in the corresponding LC-MS analysis of the tryptic digestion samples with DTT pretreatment (C and D).

>MycP1
***MQAGL**TRACQSFTAAERERSDSGVHRTLLTMVALALLTAPPALAI**DP**PSIDPGAVPPDVTGPD**Q**PT**EQ**RVLCTSP**TT**LPDS
SFHDPWNSNAYMGVGE**AHK**FATGAGVTVAVIDTGVDASPRVPAEPGGDFVDQAGDGLSDCDAHGTLTASIIIGGRPAPT**DG**
FVGVADPVRLLSLRQTSE**A**FEFVGSQPNPNPNATPAAGS**IR**SLAR**AV**VHAANLGAGVINISEAAC**Y**KVSRPIDEISLGA
AIDYAVNAKNV**V**VVAAAGNTGGDCSQNPMPDASTPNDRGW**N**KVQTVVTPAWYAPLVLTVGGIGQNGV**P**SSFSMHGPWVG
VAAPAENI**I**ALGDHGEFVNALQGR**E**GPVPIAGTS**F**AAAYV**S**GLAALV**R**QRFPELTPVQ**VM**NRITATARHPGGGIDNLVGA
GVVNAVAAL**T**WDIPPGPASVPPSV**R**RLPP**P**RIE**P**GDHRPITMVAVSVLGLTLVLGLGLTLAAR**LR**RR*

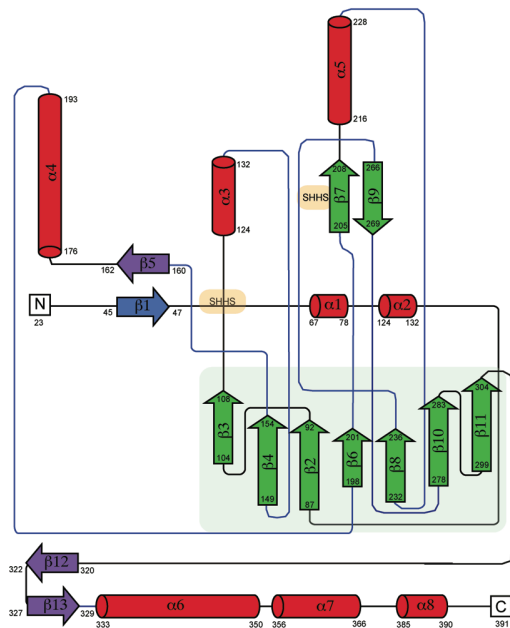
>MycP3
***M**KRAGSACLTAILAI**F**TTGALWAAPPVSAI**S**PPAVDD**S**VQPPSGT**P**GPVQ**T**ME**Q**R**G**ACSV**S**GLIP**G**TD**I**SV**P**ASQAMLN
L**P**AAWQFSR**G**EGQLVA**I**LD**T**GVQ**P**PRLPNV**D**AGGDFVDSDGLTDCDGRGTLV**A**GIVAGQ**P**NGDFSGVAP**A**ARLLS**I**R
V**T**SDKFSARSSGGDP**S**VERATIDVAALGRAIVHAAD**R**GARVINISSITCLPADRQVDQAALGAA**I**RYAAVDKDV**V**IVAAA
GNSGAGGSGGGT**C**ESNPLTDLSHPDDPRNWAGVTSV**S**IPSLWQ**P**YVLSVASL**T**AGGQ**S**KFTMAGPWVG**I**AAPGENIVS
VSNADGGGLANGLPNQ**R**Q**L**TPLSGTSYAAGYVSGVAALVRSKY**P**ELTAAQ**V**MHRLTATAHNGAR**D**PSNVV**G**AGILDPA**A**
AL**T**WELPAADGGAPQ**V**APAKPVSVPPAPQ**P**RNTTPRN**V**AFAGAAALTVLVAGVAATVAIARR**R**KDAEG**R**TAREL*

>MycP5
***M**Q**R**FGTVSGRL**R**PGRGSTATIAALLLASGALAGLPPAYAIN**P**PTIDLGALPPDG**T**PGPPAP**M**KQNSYCTEVGV**L**PG**T**DFK
LQ**F**KY**M**EMLN**I**NEAWQ**F**GRGAGVKVAVIDTGVT**P**HP**R**FP**H**LIPGGDYVMGGDGLQDCDAHG**T**IVAS**M**IGAAPANGAL**P**PP
AV**P**RR**P**VTIP**T**TEK**P**PP**P**Q**T**VLSPVPP**Q**TVTVIPG**P**PEEGAP**Q**GE**P**GP**G**VPP**P**AP**Q**PPASNHGGGT**V**IPSYSGGA
R**V**TGVDHAGGPR**L**DP**P**PPAPDAFSGIAPEVELISIRQSSQAFGLKDPY**T**GDEDP**Q**TQ**Q**KIDDVETMARAI**V**HANMGAS
V**I**NISDV**T**CMSAR**N**VIDQNALGA**A**VHYAAVDK**N**VVIVAAAGDGS**K**KDCKQ**N**PIFDPLQ**P**DDPR**D**WNA**V**TT**V**TP**S**WFSDY
V**L**TVGAVDTNG**Q**PMT**K**MSIAGPWV**S**IAAPGT**D**VIGLSPRDDGLIN**A**IDGPD**N**SLLVPAGTS**F**ATA**I**VSGVVALV**R**AK**Y**PE
LSAYQ**I**R**N**RLI**H**TAR**P**PARGVD**N**QVGYGVVD**P**VAAAL**T**WDV**P**EG**P**V**K**PK**Q**LSAP**L**EL**P**K**P**PA**E**RMV**P**VVVAAGGL**T**GAL
LIGGAVFGTATLMRR**S**R**K**RR*

Fig. S3. Mycosin-derived peptides detected by MS/MS analysis.

Amino acid sequences of MycP_{1mm}, MycP_{3mm} and MycP_{5mm}. Peptide sequences that were detected by MS/MS analysis are highlighted in grey. Potential trypsin digest sites are underlined. The signal sequences and N-terminal extensions are indicated in italic and bold fonts, respectively. The data were derived from: van der Woude *et al.* (2013) J. Bacteriol. 195, 2050-2059.

A



B

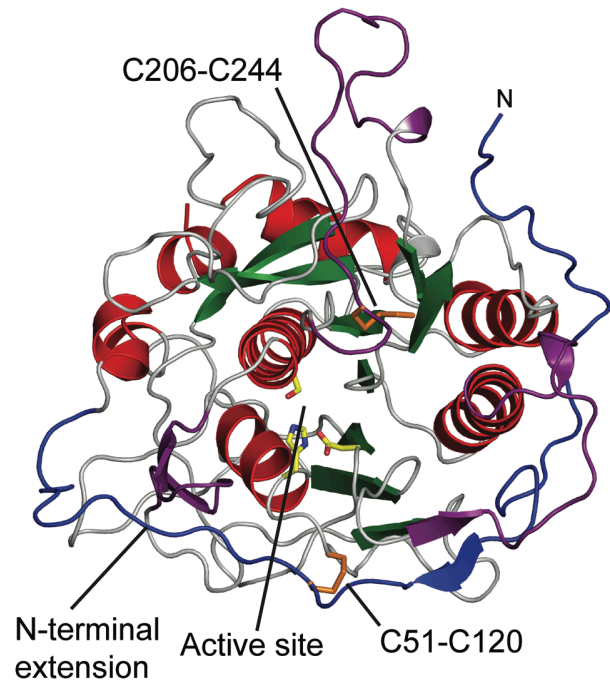


Fig. S4. Topology diagram of MycP_{1mth}.

A. MycP_{1mth} consists of a 7 stranded β sheet core (green) surrounded by 8 α helices (red). Three small anti-parallel β sheets are located surrounding the active site pocket. Two of these small β sheets include neighboring disulfide bonds (SHHS). Elements contributed by the N-terminal extension are colored blue. Elements contributed by insertion loops 1-3 are colored purple. Secondary structure diagram made with the help of the Pro-origami server (<http://munk.csse.unimelb.edu.au/pro-origami/>).

B. Cartoon of MycP_{1mth} with secondary structure elements colored as in (A). The catalytic triad is shown as yellow sticks. Side chains of cysteine residues forming disulfide bonds are shown as orange sticks.

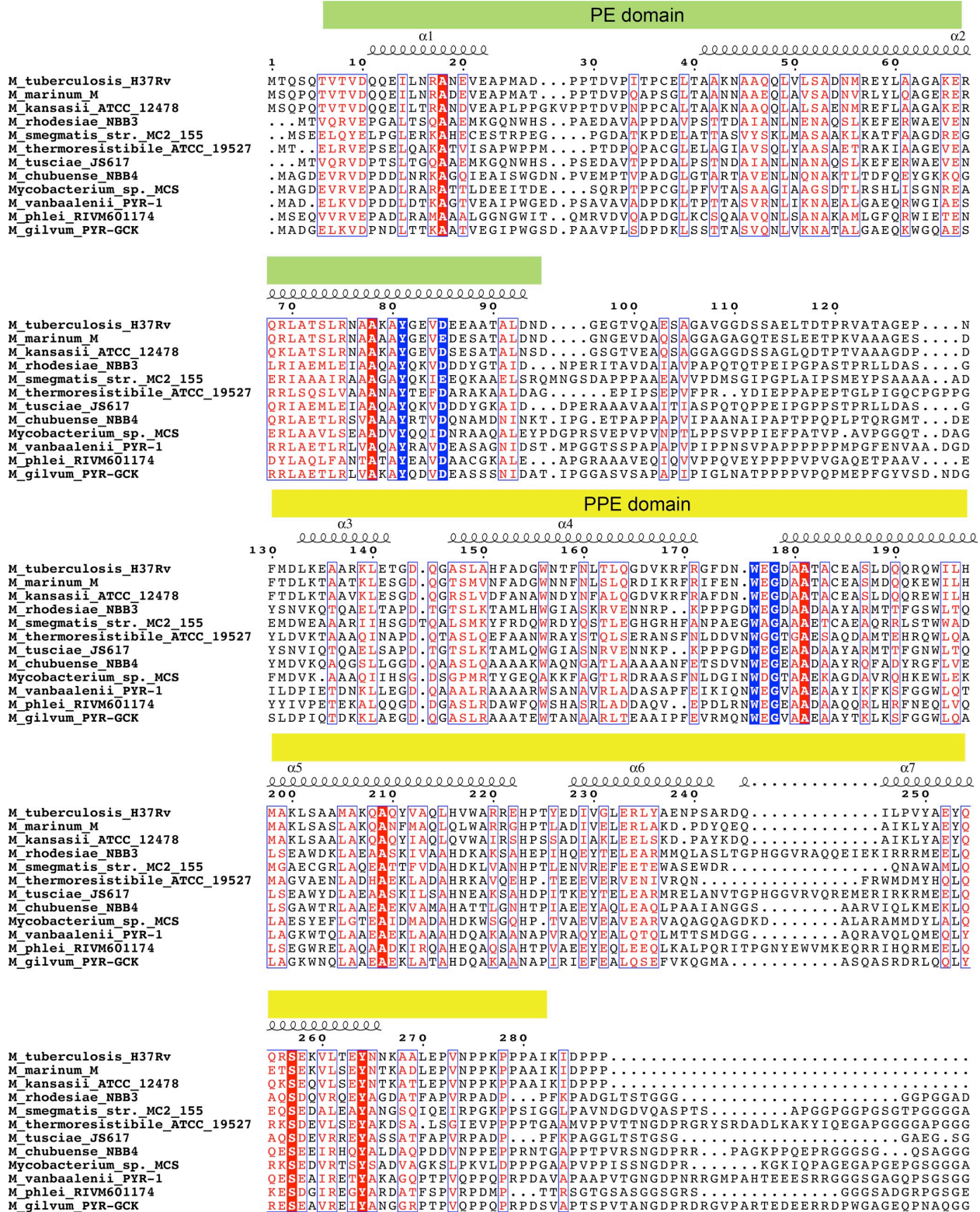


Fig. S5. Sequence alignment of EspB proteins (continues on the next page).

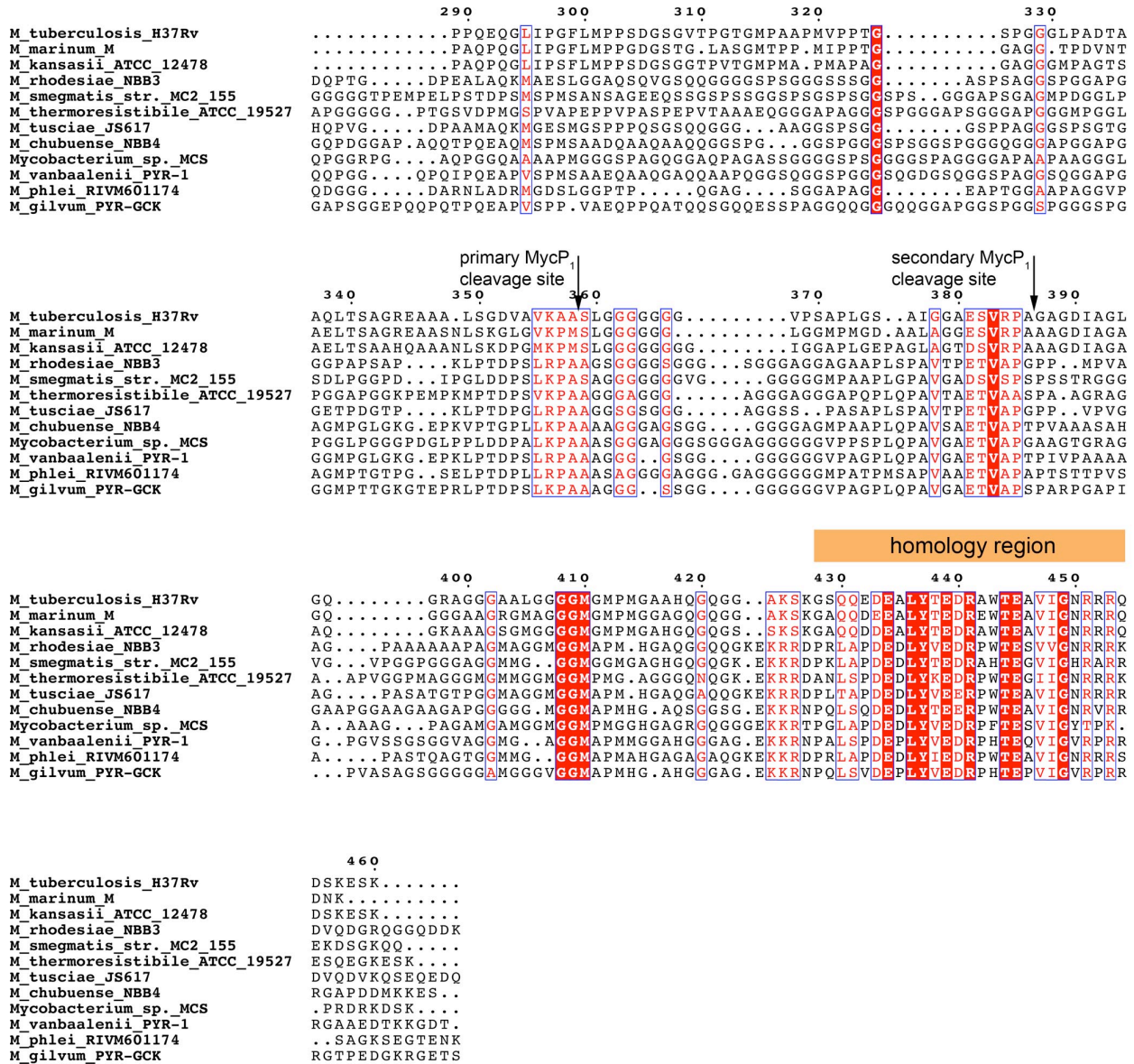


Fig. S5. Sequence alignment of EspB proteins.

The amino acid sequence alignment of EspB proteins from the ESX-1 clusters of mycobacteria was prepared using ClustalW2 (<http://www.ebi.ac.uk/Tools/msa/clustalw2/>). The PE and PPE domains, and the secondary structure elements are indicated as predicted by the HHpred server (<http://toolkit.tuebingen.mpg.de/hhpred>). The MycP₁ cleavage sites are indicated by vertical arrows. The YxxxD/E type VII (ESX) secretion motif in PE domain and the WxG motif in PPE domain are highlighted in blue. EspB proteins have a homology region of unknown function, indicated by orange rectangle.