Supplementary Information

Experimental Identification and Computational Characterization of a Novel Extracellular Metalloproteinase Produced by *Clostridium sordellii*

Michael J. Aldape[1,2,*]; Aoxiang Tao[3]; Dustin D. Heeney[1]; Eric R. McIndoo[1]; John M. French[1]; Dong Xu[3,*]

[1]Veterans Affairs Medical Center, Boise, ID, 83702, USA; [2] Northwest Nazarene University, Nampa, ID, 83686, USA; [3] Department of Biomedical and Pharmaceutical Sciences, College of Pharmacy, Idaho State University, Meridian, ID, 83642, USA

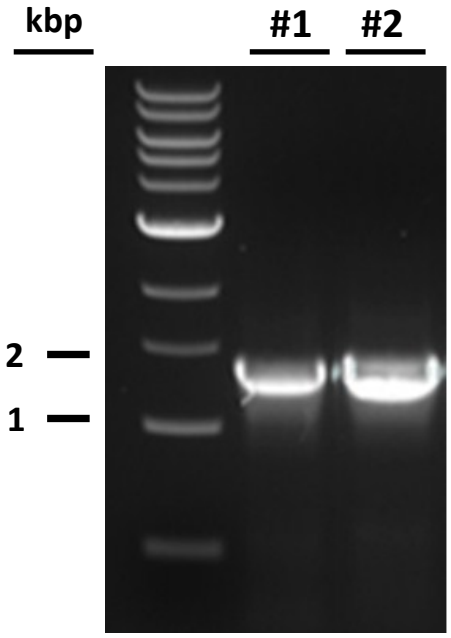
*Co-Corresponding authors:

- Michael John Aldape, Veterans Affairs Medical Center, Research and Development, Infectious Diseases Section, 500 West Fort St. (Bldg 117), Boise, ID 83702. Phone: 208-422-1000 x7659; Fax: 208-422-1425; e-mail: mike.aldape@va.gov
- Dong Xu, Department of Biomedical and Pharmaceutical Sciences, College of Pharmacy, Idaho State University, Meridian, ID, 83642. Phone: 208-373-1832; Fax: 208-373-1834; e-mail: dxu@pharmacy.isu.edu

Supplementary Figures:

Figure S1:

(A)



(B)

Strain	Locus Tag	ORF length (kb pairs)
ATCC 9714	9714_00259	1,538
630	JGS6382_01930	1,538
UMC 164	UMC_00586	1,538
DA-108	DA108_00265	1,538

Figure S2:

(A)

atgattaaaaaaaatatagcaattttaacagtagcaggattaactctcagttttctagcaccttcagtaagttatgcagatagtcaaacaaaagaa
gatttaccacaagcatataatgttaaatcagaaaaagatttaaaggatgttttgaaaaatgtatcagatgaatataatataaaaaagaatactgat
caatatcaattagttgatacagaaaaagacgatttaggatttactcattatactttaaaacctaaagctgatggctattttgctgaagcagcagaa
gttaaagtacatactgataagtctggcaaagttgtctttgtaaatggagatttagatcaaggtaaattggaagttaaaaatcaaactaaaatagat
aaagataaagctattgaattagcttttaaagcaatagaaaaatcacgtgacgaagttaaaaatttatcaggtaaagatgttgtacaagaagctaa
aatagccgtagatgagcaaagtaatagagcaatttatagcttagacatagcgtatactgtaccagagccagctcattggaatattaagttagat
gcagaaaatggaaatataattgaaaaacaaaatattatagaaaatgcagctcaaacaacaggaactggaataggtgccaatggacaagtta
aatcacctttaaatataacagaagatggaggaaagttctatttattagatactactcataaagggaaaatagctactatagcttttgaggcatttg
atgataatggtatagtaggaagtgttatatctaatataaaaaattgttttgatggtgaaaaagataaagcagccgtagatgctcactactttacag
ataaagtttataaatattataaagatgttcataatagagaaagttatgatgacaaaggatcagatatatactcttatgtacatgtaccaaatccaga
tacaggtgaaagttggagtaatgctgcttggactggagcagagatgatttatggagatggtaatcaagttgaagaaaattctttctcagcagca
gatgatgttgttgctcatgaaattacacatggagttacatcatcaacagctaatttagtttataagtatcaaccaggagctttaaatgaatctttttc
agatgtatttggatactttgtagatagtgatgattggacaatgggtgaagatttatataaaacacctaatacagctataagagatttaaaagaacc
taaaaaatataatcaaccagaacatatgaatgaatataaaaattatagtataaattatgatagaggtggagttcatataaatagtggtattcctaa
caaagcagcatataacactataactaaattaggaaaagaaaaagctgaaaaaatttactatagagctcttactaaatatttaacaagacaatca
gacttctcagatgctagtcatgctttaactcaatcagcaaaagatctttatggagataaagaagctaatgatgttcaagctgcatggaatgaggt
tggagtaaattaa

(B)

MIKKNIAILTVAGLTLSFLAPSVSYADSQTKEDLPQAYNVKSEKDLKDVLKNVSDEYN

IKKNTDQYQLVDTEKDDLGFTHYTLKPKADGYFAEAAEVKVHTDKSGKVVFVNGDLD

QGKLEVKNQTKIDKDKAIELAFKAIEKSRDEVKNLSGKDVVQEAKIAVDEQSNRAIYSL

DIAYTVPEPAHWNIKLDAENGNIIEKQNIIENAAQTTGTGIGANGQVKSPLNITEDGGKFY

LLDTTHKGKIATIAFEAFDDNGIVGSVISNIKNCFDGEKDKAAVDAHYFTDKVYKYYKD

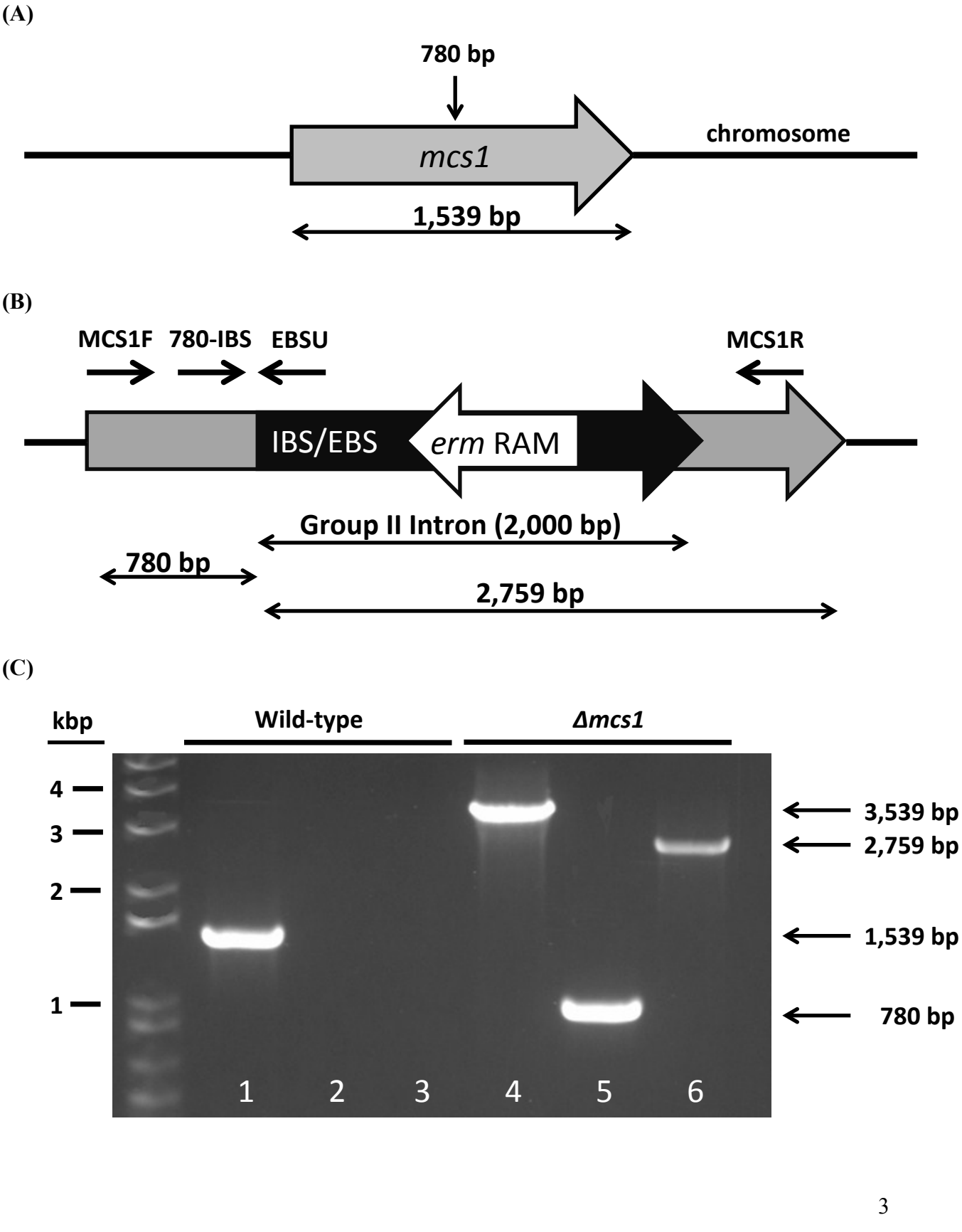
VHNRESYDDKGSDIYSYVHVPNPDTGESWSNAAWTGAEMIYGDGNQVEENSFSAADD

VVA**HEITH**GVTSSTANLVYKYQPGALNESFSDVFGYFVDSDDWTMGEDLYKTPNTAIR

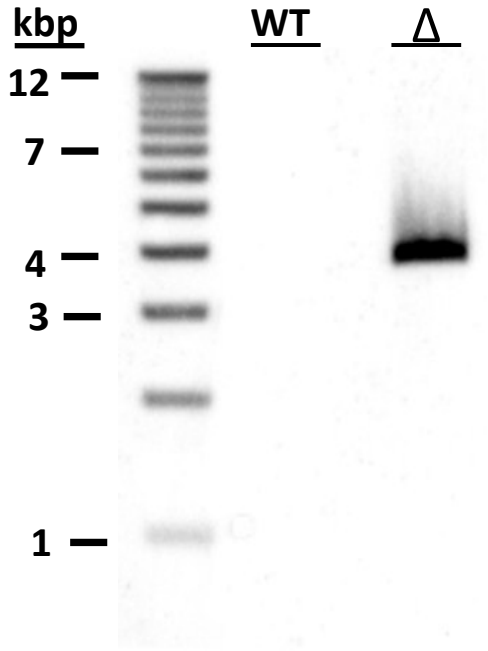
DLKEPKKYNQPEHMNEYKNYSINYDRGGVHINSGIPNKAAYNTITKLGKEKAEKIYYRA

LTKYLTRQSDFSDASHALTQSAKDLYGDKEANDVQAAWNEVGVN

Figure S3.


(A)
780 bp
chromosome
mcs1
1,539 bp
(B)
MCS1F
780-IBS
EBSU
MCS1R
IBS/EBS
erm RAM
Group II Intron (2,000 bp)
780 bp
2,759 bp
(C)
kbp
Wild-type
Δmcs1
4
3
2
1
3,539 bp
2,759 bp
1,539 bp
780 bp
1
2
3
4
5
6

(D)



Supplementary Figure Legends:

Figure S1. Identification of the Mcs1 open reading frame in multiple *C. sordellii* strains. The Mcs1 open reading frame was identified in A) in 2 out of 2 additional clinical isolates using PCR analysis, and B) 3 out of 3 *C. sordellii* strains following annotated whole genome sequence examination.

Figure S2. Nucleotide and protein sequences of the *C. sordellii* protease Mcs1. A) Nucleotide sequence of the Mcs1 (ORF_00259) open reading frame. B) The deduced amino acid sequence of the pro-peptide of Mcs1. The 512-aa peptide contains a 26 amino acid leader sequence (bolded and underlined in black) and the highly conserved zinc-binding motif sequence HEXXH (bolded and underlined in red).

Figure S3. Screening a *C. sordellii mcs1* mutant candidate by PCR. A and B) Schematic representation of A) wild-type and B) mutated *mcs1* open reading frame. The ClosTron insertion site is depicted by the vertical arrow at the base pair 780. The locations of *mcs1* specific primers, MCS1F and MCS1R, and intron specific primers, 780-IBS and EBSU, are shown with horizontal arrows on either side of the mcs1 gene and ClosTron insertion site, respectively. The sizes of the PCR products are indicated. C) PCR products of the wild-type (Lanes 1 to 3) and *mcs1* mutant (Lanes 4 to 6) strains. Lanes 1 and 3; primer pair MCS1F and MCS1R. Lanes 2 and 5; primer pair MCS1F and EBS-U. Lanes 3 and 6; primer pair 780-ISB and MCS1R. (D) Southern blot of *mcs1* mutant in *C. sordellii* ATCC 9714. Genomic DNA was harvested from wild-type (Lane 1) and Δ*mcs1* mutant (Lane 2) strains from overnight cultures. Chromosomal preparations were cleaved with PacI restriction enzyme, separated on a 0.8% agarose gel and transferred to a nylon

membrane. The membrane was hybridized with a biotinylated DNA probe spanning the *erm* RAM and 3' region of the group II intron.