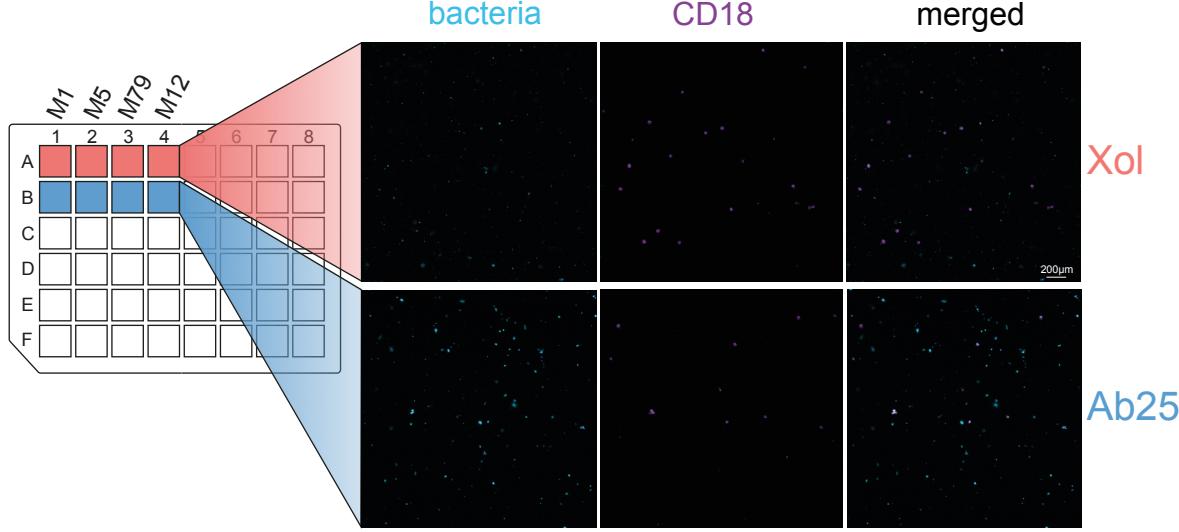


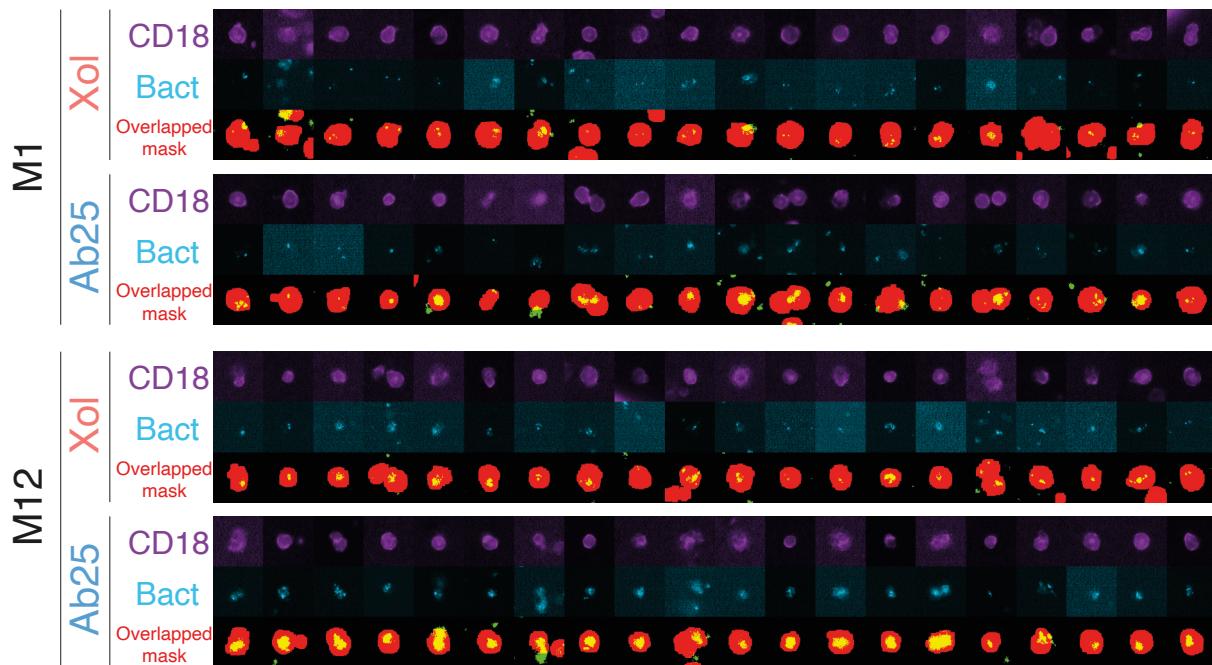
Table of contents - Appendix

- Appendix Figure S1. Imaging-based whole blood phagocytosis.
- Appendix Figure S2. DDA peptide spectra
- Appendix Figure S3. SDS-PAGE analysis of the antibodies used for the binding assays.
- Appendix Table S1. Statistical test p values.
- Appendix Data 1. M protein amino acid sequences.
- Appendix Data 2. M protein alignments.

A

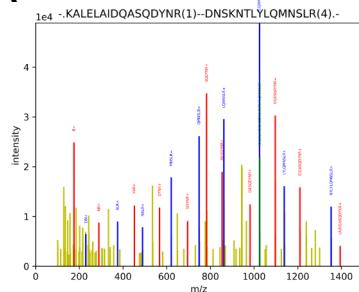
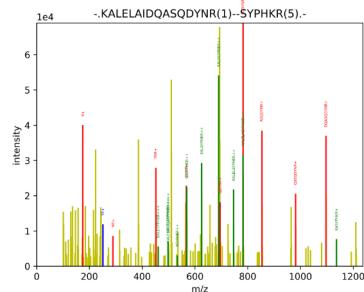
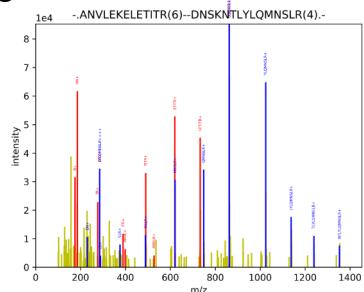
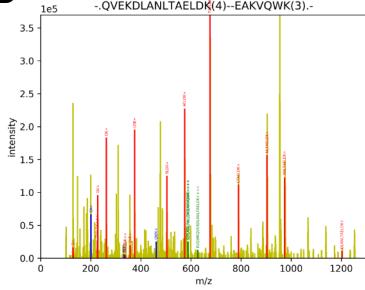
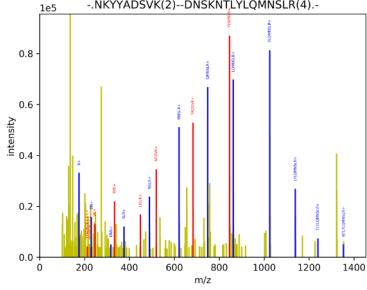
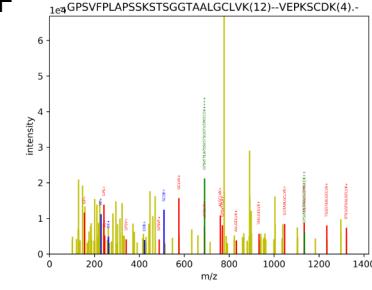
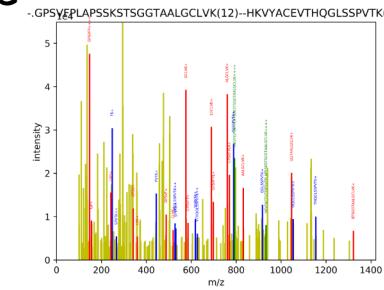
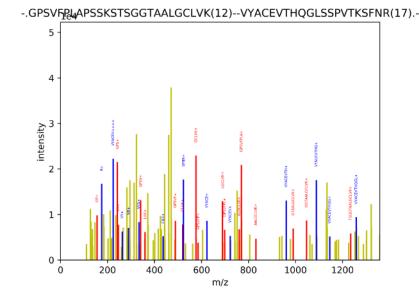
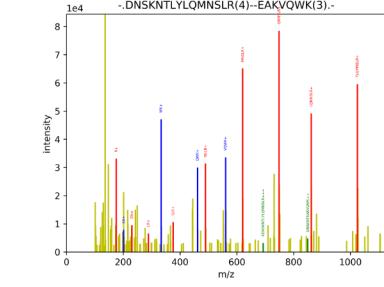
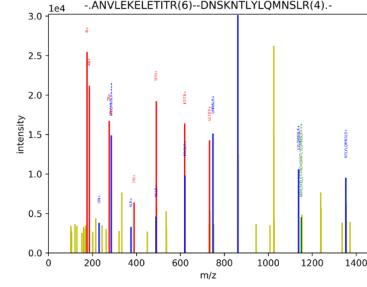
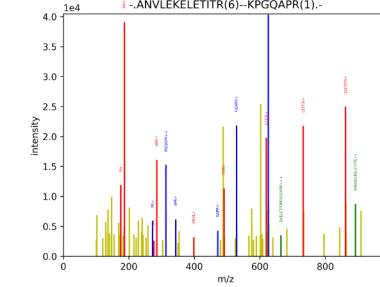


B



Appendix Figure S1. Imaging-based whole blood phagocytosis.

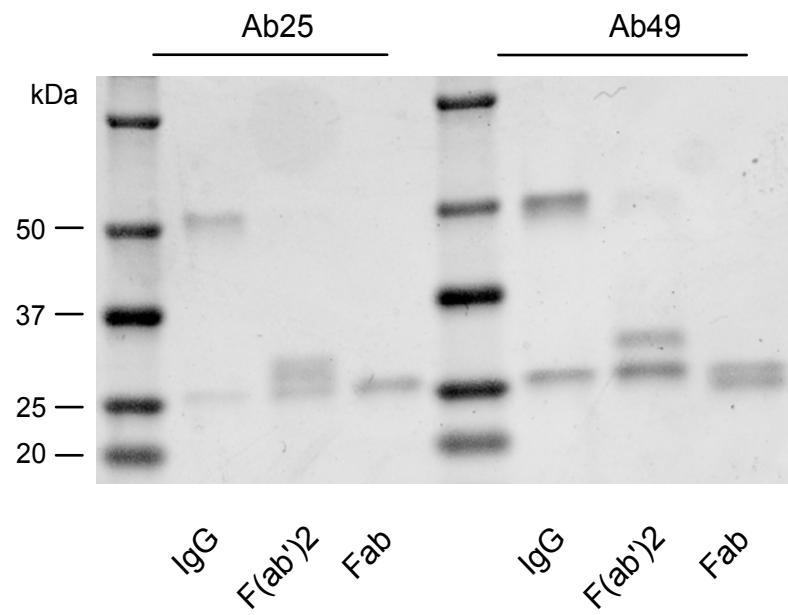
Imaging-based whole blood phagocytosis. A) Whole blood was diluted and infected with 4 different strains of GAS expressing GFP (FITC). After 30 minutes, the infection was fixed with PFA and the cells were stained with anti-CD18 (BV421, shown in magenta). The blood was immobilized onto glass-bottom 96-well plates which were precoated with anti-CD29. Images were acquired on a widefield epifluorescence microscope using a 20X objective (NA = 0.75). The acquisition of images was automated using Nikon JObS by scanning through a 96 well plate acquiring 25 images/sample. GFP (bacteria) and BV421 (CD18 positive leukocytes) were segmented based on a signal above 5 and 4 std of the background respectively. The overlapping mask for both channels was generated and the area of overlap was measured. Contrast adjustment is set to the same for all shown images. More than 560 cells were analyzed for this experiment B) Magnified images of cells representative cells were randomly selected and compared side-by-side to highlight the GFP area change with and without pre-opsonization. The presented images are chosen from a dataset representing the median of each population. The differences seen across the two groups, therefore, are fully representative of the data shown in Fig 3F.

A**B****C****D****E****F****G****H****J****K****L**

Appendix Figure S2. DDA peptide spectra

Example DDA peptide spectra of cross-linked interactions identified (see also Table EV 1 for more details). One spectrum is shown for each cross-linked peptide pair. The red and the blue peptides represent parent peptides, respectively, whereas the green signal arises from the aforementioned peptides containing an additional DSS cross-linker arm.

Antibody fragment generation



Appendix Figure S3. SDS-PAGE analysis of the antibodies used for the binding analysis

Appendix Table S1 - p values

Comparison **p value**

Fig 2

B

Ab25 vs. Xolair	0.0412
Ab49 vs. Xolair	0.0126
E	
M1+xol vs. M1+Ab mix	0.0001
M1+xol vs. M1+Ab25	0.0421

Fig 3

B

Xolair vs. IVIG	0.0556
Xolair vs. Ab25	0.0032

D

Xolair vs. IVIG	0.0292
Xolair vs. Ab25	0.0014

F

M1 Xol P1 vs. M1 Ab25 P1	<0,0001
M1 Xol P1 vs. M1 Ab25 P2	<0,0001
M1 Xol P1 vs. M1 Ab25 P3	0.9995
M5 Xol P1 vs. M5 Ab25 P1	0.8371
M5 Xol P2 vs. M5 Ab25 P2	>0,9999
M5 Xol P3 vs. M5 Ab25 P3	>0,9999
M12 Xol P1 vs. M12 Ab25 P1	<0,0001
M12 Xol P2 vs. M12 Ab25 P2	<0,0001
M12 Xol P3 vs. M12 Ab25 P3	<0,0001
M79 Xol P1 vs. M79 Ab25 P1	<0,0001
M79 Xol P2 vs. M79 Ab25 P2	0.9988
M79 Xol P3 vs. M79 Ab25 P3	0.0322

E

Ab25 vs Xolair	<0,0001
----------------	---------

Fig 4

Liver

IVIG vs Untreated	0.0011
Ab25 vs untreated	0.0027

Kidney

IVIG vs Untreated	0.0329
Ab25 vs untreated	0.0024

Spleen

IVIG vs Untreated	0.0329
Ab25 vs untreated	0.0043

Fig 6

A

Untreated vs. Ab25	0.0005
Untreated vs. Ab32	0.9607
Untreated vs. Ab49	0.4265
Untreated vs. IVIG 10ug/ml	0.6491
Untreated vs. IVIG 100ug/ml	0.8049

B

Xolair vs. Ab25	<0,0001
Xolair vs. Ab25 fabs	0.0784

C

Ab25 vs. Ab25 Ides	>0,9999
Ab25 vs. Ab25 fabs	<0,0001
Ab25 Ides vs. Ab25 fabs	<0,0001
Ab49 Ides vs. Ab49 fabs	<0,0001
Ab49 vs. Ab49 Ides	0.3676
Ab49 vs. Ab49 fabs	<0,0001

D

Plasma vs. Plasma fabs	<0,0001
Plasma Ides vs. Plasma fabs	<0,0001

F

M1	
Ab25 IgG vs. Ab25 Fabs	<0,0001
Ab25 F(ab') ₂ vs. Ab25 Fabs	0.0001

M5

Ab25 IgG vs. Ab25 Fabs	0.0031
Ab25 F(ab') ₂ vs. Ab25 Fabs	0.0071

M12

Ab25 IgG vs. Ab25 Fabs	0.005
Ab25 F(ab') ₂ vs. Ab25 Fabs	0.2286

M79

Ab25 IgG vs. Ab25 Fabs	<0,0001
Ab25 F(ab') ₂ vs. Ab25 Fabs	<0,0001

Fig EV 2

Xolair vs. Ab25	0.0001
Xolair vs. Ab32	0.0002
Xolair vs. Ab49	0.0002
Xolair vs. Plasma	<0,0001

Data EV1. M protein sequences

M1-90-226-CTG	>gene_1147 GeneMark.hmm 306_aa + 113293 114210 >NODE_5_length_114211_cov_48.735082 MAKNNTNRHYSRKLKLTGTASVAVALTVLGAGFANQTEVKANGDGNPREIEDLAANNPA IQNIRLRHENKDLKARLENAMESVAGRDFKRAEELEKAKQAALEDQRKDLETKLKELQQDYD LAKESTSWDRQRLEKELEEKKEALELAIDQASRDYHRATALEKELEEKKALELAIDQAS QDYNRANVLEKELETITREQEINRNLLGNAKLELDQLSSEKEQLTIEKAKLEEKQISDA SRQSLRRDLDASREAKKQVEKDLANLTAELDKVKEEKQISDASRQGLRRDLDASREAKKQ VEKDLA
M4-BS4-CTG	>gene_436 GeneMark.hmm 222_aa + 185954 186619 >NODE_2_length_186620_cov_148.285713 MARKDTNKQYSRKLKLTGTASVAVAVAVLGAGFANQTEVKAAEIKKPQADSAWNWPKEYN ALLKENEELKVEREKYLSYADDKEKDPQYRALMGENQDLRKREGQYQDKIEELEKERKEK QERQEQLERQYQIEADKHYQEQQKKHQEQQQLEAKQKLAQDKQISDASRQGLSRDLEA SREAKKKVEADLAALTAEHQKLKEDKQISDASRQGLSRDLEA
M5-Manfredo-CTG	>gene_1398 GeneMark.hmm 194_aa + 61232 61813 >NODE_10_length_61815_cov_235.110363 MARENTNKHYSRKLKKGTTASVAVALSVLGAGLVVNTNEVSAAVTRGTINDPQRAKEALD KYELENHDLTKNEGLKTENEGLKTENEGLKTENEGLKTKEHEAENDKLKQQRDTLST QKETLEREVQNTQYNNETLKIKNGDLTKELNKTRQELANKQQESKENEKALNELLEKTVK DKIAKEQENKETIG
M8-BB3-CTG	>gene_1160 GeneMark.hmm 238_aa - 2 715 >NODE_7_length_101418_cov_178.048943 MARKDTNKQYSRKLKLTGTASVAVAVAVLGAGFANQTEVKAEASKHSISNNEQLINLN DLIEENNNDLKDQLARNLDDNTREKDPQYRALMGENQDLREKEGKYQDKIKKLEEKEKN LEKKSEDVERHYLKKLDQEHKEQQRQKNLEELERQSQRREIDKRYQEQLQKQQLETEKQ ISEASRKSLSRDLEASRAAKDLEAEHQKLKEEKQISDASRQGLSRDLEASREAKKV
M12-BS3-CTG	>gene_1710 GeneMark.hmm 575_aa + 13619 15346 >NODE_21_length_15967_cov_54.040096 MAKNTTNRHYSRKLKLTGTASVAVALTVVAGLVAGQTVRADHSDLVAEKQRLEDLGQKF ERLKQRSELYLQQYYDNKSNGYKGDWYVQQQLKMLNRDLEQAYNELSGEAHKDALGKLGD NADLKAKITELEKSVEEKNDVLSQIKKELEEAKEKDIOFGREVHAADLLRHKQEIACEKENV ISKLNQELQPLKQKVDETDRNLQQEKQKVLSEQQLAVTKENAKKDFFELAALGHQLADKE YNAKIAELESKLADAKDFELAALGHQAHNEYQAKLAEKDGQIKQLEEQKQILDASRKG TARDLEAVRQAKKATEAEELNNLKAELAKVTQKQILDASRKGTTARDLEAVRKAKAQVEAA LKQLEEQNPKISEASRKGLRRDLDASREAKKQVEKDLANLTAELDKVKEEKQISDASRQGL RRDLDASREAKKQVEKALEEANSKLAALEKLNKELEESKKLTEKEKAELQAKLEAEAKAL KEQLAKQAEELAKLRAKGASDSQTPDAKPGNKAVPGKGQAPQAGTKPNQNPKAPMKTKRQ LPSTGEAAPFFTAAALTVMATAGVAAVVKRKEEN
M28-BB5-CTG	>gene_950 GeneMark.hmm 358_aa - 4735 5811 >NODE_5_length_210162_cov_207.134746 MARKDTNKQYSRKLKLTGTASVAVAVAVLGAGFANQTEVKAAEASKSTETSANGADKLAD AYNTLLTEHEKLRDEYYTLIDAKEEPRYKALRGENQDLREKEGKYQDKIKKLEEKEKNL EKKSEDVERHYLKKLDQEHKEQERQKNLEELERQSQRREIDKRYQEQLQKQQLETEKQI SEASRKSLSRDLEASRAAKDLEAEHQKLKEEKQISDASRQGLSRDLEASREAKKVEAD LAEANSKLALEKLNKELEEGKKSEKEKAELQARLEAEAKALKEQLAKQAEELAKLKG QTPNAKVAPQANRSRSAMTQQKRTLPSTGEAAPFFTAAAATVMVSAGMLALKRKEEN
M75-BS7-CTG	>gene_561 GeneMark.hmm 387_aa - 31358 32521 >NODE_3_length_208685_cov_62.045909 MARKDTNKQYSRKLKLTGTASVAVAVAVLGAGFANQTEVKAAEERTFTELPYEARLYKAWK SENDELRENYRRTLDKFNTEQGKTRLEEQNKKLHSELASVTETLTSVTEADDKKIKDLT DRDKISSNLLGNAKDQINKLTTEKDTLAEKAKKLEEDKQISDASRKSLSRDLEASRAAKK ELEANHQKLETEHKKLKEEKQISDASRQGLSRDLEASREAKKVEADLAALTAEHQKLKE EKQISDASRQGLSRDLEASREAKKVEADLAEEANSKLALEKLNKELEEGKKSEKEKA LQARLEAEAKALKEQLAKQAEELAKLKGNTQTPNAKVAPQANRSRSAMTQQKRTLPSTGEA ANPFFTAAAATVMVSAGMLALKRKEEN

M79-BTS4-CTG	>gene_435 GeneMark.hmm 222_aa - 185954 186619 >NODE_2_length_186620_cov_58.625015 MARKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTEVKAAEIKKPQADSAWNWPKEYN ALLKENEEALKVEREKYLSYADDKEKDPQYRALMGENQDLRKREGQYQDKIEELEKERKEK QERQEQLERQYQIEADKHYZEQQQKHHQQEQQQLEAEKQKLAQDKQISDASRQGLSRDLEA SREAKKKVEADLAALTAEHQKLKEDKQISDASRQGLSRDLEA
M81-BB6-CTG	>gene_94 GeneMark.hmm 396_aa - 101371 102561 >NODE_1_length_190029_cov_107.359270 MVRKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTEVKAAGSEENVPKQQYNALWEEN EDLGRGRERKYIAKLEKEEIQNGELNEKNRKLEADIADLQDVIEDNDQEIKRKDRMYEAFL KQSKDQVKDLTAEKDTLAEKAKKLEEDKQISDASRKSLSRDLEGSRAAKKELEAKHQKLE TEHQKLKEDKQISDASRQGLSRDLEASREAKKKVEADLAALTAEHQKLKEEKQISDASRQ GLSRDLEASREAKKKVEADLAEANSKLQALEKLNKELEEGKKLSEKEKAELQARLEAEAK ALKEQLAKQAEELAKLРАГKASDSQTPDAKPGNКVVPKGQAPQAGTKPNQNКАPMKETK RQLPSTGEAANPFFTAAAATVMVSAGMLALKRKEEN
M85-SP5-CTG	>gene_158 GeneMark.hmm 402_aa - 835 2043 >NODE_2_length_119743_cov_614.138920 MVRKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTEVKAAEQRAPQAKGTSVSADLYN SLWDENKTLREKQGEYITKIQNEETKNKELDKKNKELDNSRTVDLIDVIEHDDQELERKER MYEAFLKQSKDQVNNLTAEKDTLAEKAKKLEEDKQISDASRKSLSRDLEGSRAAKKELEA KHQKLETEHQKLKEDKQISDASRQGLSRDLEASREAKKKVEADLAALTAEHQKLKEEKQI SDASRQGLSRDLEASREAKKKVEADLAEANSKLQALEKLNKELEEGKKLSEKEKAELQAR LEAEAKALKEQLAKQAEELAKLРАГKASDSQTPDAKPGNКVVPKGQAPQAGTKPNQNKA PMKETKRQLPSTGEAANPFFTAAAATVMVSAGMLALKRKEEN
M87-BB10-CTG	>gene_1714 GeneMark.hmm 238_aa - 3 716 >NODE_26_length_15865_cov_62.755661 MARKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTEVKAESPREVTNELAASVWKKV EEAKEKASKLEKQLEEAQKDYSEIEGKLEQFWHDYDKLEKENKEYASQLGNQEEREKLE LEYLRKSDEEYKEHQQRQEERQKNLEELERQNKRIDLKRYQEQLQKQQQLETEKQIS EASRKSLSRDLEASRAAKKELEAEHQKLKEEKQISDASRKSLSRDLEASREAKKKVEA
M89-BSA4-CTG	>gene_1171 GeneMark.hmm 244_aa - 2 733 >NODE_8_length_78604_cov_108.090657 MARKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTTVKADSDNINRSVSVDNEKELH NIKIADLEERGEHLDKIDELKEELKAKEKSSENVERHYLRKLDQEYKEQQERQKNLEELE RQSQREVEKRYQEQLQKQQQLETEKQISEASRKSLSRDLEASRAAKKDEAEHQKLKEEK QISDASRQGLSRDLEASREAKKKVEADLAALTAEHQKLKEEKQISDASRQGLSRDLEASR EAKK
M118-SP1-CTG	>gene_1703 GeneMark.hmm 332_aa - 4820 5818 >NODE_14_length_20826_cov_45.853546 MARKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTEVKAAEKKVEVADSNASSVAKLY NQIADLTDKNGEYLERIEELEERQKNLEKLERQSVAADKHYZEQVKKHQEYKQEERQ KNLEELERQNKRIDLKRYQEQLQKQQQLETEKQISEASRKSLSRDLEASRAAKKDEAEH QLKKEEKQISDASRQGLSRDLEASREAKKKVEADLAEANSKLQALEKLNKELEEGKKLSE KEKAELQAKLEAEAKALKEQLAKQAEELAKLKGНQTPNAKVPQANRSRSAMTQQKRTL STGETANPFFTAAAATVMVSAGMLALKRKEEN
M179-BB12-CTG	>gene_92 GeneMark.hmm 433_aa - 100604 101905 >NODE_1_length_264627_cov_344.813212 MVRKDTNRHSYSLRKLKTGTASVAVAVAVLGAGLAVNQTEVSAKSVTRSTAQDPDKSRQAI TEYEVENHKLTQEKNALTNRNQELTDENGELKTANEALRQRGDTLFNQRVKEKQVQEKE HNNKTLKIENGELKTENGDLTKLDETRQELANKQQESKENKTLNELLEKTVKDIAKE QENKETIGTLKLLDETVKDKIAKEQKSKQDFGALKQELAKKEEQNQISDASRQGLRRDL NASREAKKKVKEKQVDSLARQGLRRDLDASREAKKKVEKALEANSKLAALEKLNKELEES SKLALEKLNKELEESKKLDEKEKAELQAKLEAEAKALKEKLAKQAEELAKLРАГKASDS QTPDAKPGNКVVPKGQAPQAGTKPNQNКАPMKETKRQLPSTGEAANPFFTAAAATVM AGVAAVVKRKKEN

Data EV2. Sequence similarity and alignment analysis

The genome sequencing data were searched against the CDC database of M protein families to detect the target M protein sequence. Each M protein sequence was pairwise aligned with the target M1 protein using EMBOSS Needle web server and the result is reported below.

```
#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210916-185110-0458-22637843-p1m.asequence
#   -bsequence emboss_needle-I20210916-185110-0458-22637843-p1m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
=====
# Aligned_sequences: 2
# 1: Reference_M1_sequence
# 2: M4_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
# Length: 486
# Identity: 125/486 (25.7%)
# Similarity: 157/486 (32.3%)
# Gaps: 266/486 (54.7%)
# Score: 461.0
=====

Reference_M1_ 1 MAKNNNTNRHYSLRKLKTGTASVAVALTVLGAGFANQTEVKANGDGNPREV 50
||:..||.:|||||:|||||:|||||:|||||:|||||:....|:
M4_sequence 1 MARKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTEVKAAEIKKPQ-- 48

Reference_M1_ 51 IEDLAANNPAIQNIRLRYENKDLKARLENAMEVAGRDFKRAEELEKAKQ- 99
.||.||.||...||.: ||::||...||.:|| |::||..|
M4_sequence 49 -ADSAWNWPKEYNALLK-ENEELKVEREKYLSY-----ADDKEKDPOY 89

Reference_M1_ 100 -ALEDQRKDLETKLKELQQDYDLAKESTSWSRDRQRLEKEEEKKEALELAI 148
||...|||...||...|... :|||..:|||..|
M4_sequence 90 RALMGENQDLRKREGQYQDKI-----EELEKERKEKQERQE--- 125

Reference_M1_ 149 DQASRDYHRATALEKELEKKKALELAIDQASQDYNRANVLEKELETITR 198
|...|.||. ...:||...:|| :|:
M4_sequence 126 -QLEROYQ--IEADKHYZEQQK-----K-----K 145

Reference_M1_ 199 EQEINRNLLGNAKLELDQLSSEKEQLTIEKAKLEEEKQISDASRQSLRRD 248
.|:| |::||...||.||.:|||...||.|||:
M4_sequence 146 HQQ-----EQQQLEAEKQKLAKDKQISDASRQGLSRD 177

Reference_M1_ 249 LDASREAKKQVEKDLANLTAELDKVKEDKQISDASRQGLRRDLDASREAK 298
|:|||||:||.||.||.||:|||:|||||:||.|||:
M4_sequence 178 LEASREAKKQVEADLAALTAEHQKLKDQISDASRQGLSRDLEA---- 222

Reference_M1_ 299 KQVEKDLANLTAELDKVKEEKQISDASRQGLRRDLDASREAKKQVEKALE 348
M4_sequence 223 ----- 222

Reference_M1_ 349 EANSKLAALEKLNKELEESKKLTEKEKAELQAKLEAEAKALKEQLAKQAE 398
M4_sequence 223 ----- 222

Reference_M1_ 399 ELAKLRAKGASDSQTPDTKPGNKAVPGKGQAPQAGTKPNQNPKPMKETKR 448
M4_sequence 223 ----- 222

Reference_M1_ 449 QLPSTGETANPFFTAAALTVMATAGVAAVVKRKEEN 484
M4_sequence 223 ----- 222

#-----
#-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210916-185742-0233-9205053-p2m.asequence
#   -bsequence emboss_needle-I20210916-185742-0233-9205053-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####
=====

# Aligned_sequences: 2
# 1: Reference_M1_sequence
# 2: M5_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 495
# Identity:    79/495 (16.0%)
# Similarity:  110/495 (22.2%)
# Gaps:        312/495 (63.0%)
# Score:      225.5
=====

Reference_M1_      1 MAKNNNTNRHYSLRKLKGTGATASVAVALTGLGAGF-ANQTEVKANGDGNPRE
                   ||::|||:|||||||.|||||||:|||||. .|||.| . .|.
M5_sequence       1 MARENTNKHYSLRKLKKGATASVAVALSVLGAGLVLVNTNEVSA---AVTRG

Reference_M1_      50 VIEDLAANNPAIQNIRLRYENKDLKARLENAMAEVAGRDFKRAEELAKAQ
                   .||. ....|:||| . ||.|||.:|... . .|||.:|:|:|
M5_sequence       48 TINDPQRAKEALDKYEL--ENHDLTKNEGLK-----TENEGLKTENE

Reference_M1_      100 ALEDQRKDLETKLKELOQDYDLAKESTS WDRQRLEKELEEKKEALELAID
                   .||:....|:|||.:|||.:|.:| . || ..|..|||.:|...|:|
M5_sequence       89 GLKTENEGLKTKEKHEAENDKLQ-----QR--DTLSTQKETLEREVQ

Reference_M1_      150 QASRDYHRATA----LEKELEEKKKALELA-IDQASQDYNRA--NVLE
                   :|||...| . .|||.:|:| .|| ..|..|:|:| .||| .|:|:|
M5_sequence       131 --NTQYNNETLKIKNQGDLTKELNKTRQ--ELANKQQESKENEKALNELLE

Reference_M1_      191 KEL-ETITREQEINRNLLGNAKLELDQLSSEKEQLTIEKAKLEEAKQISD
                   |.. :|.:| | | .|:|:|:|
M5_sequence       177 KTVKDKIAKEQE--NKETIG-----

Reference_M1_      240 ASRQLSLRRDLDASREAKKQVEKDLANLTAELDKVKEKDQKISDASRQLRR
M5_sequence       195 -----
Reference_M1_      290 DLDASREAKKQVEKDLANLTAELDKVKEEKQISDASRQLRRDLDASREA
M5_sequence       195 -----
Reference_M1_      340 KKQVEKALEEANSKLAALEKLNKELEESKKLTEKEKAELQAKLEAEAKAL
M5_sequence       195 -----
Reference_M1_      390 KEQLAKQAEELAKLRAGKASDSQT PDTKPGNKAVPGKGQAPQAGTKPNQN
M5_sequence       195 -----
Reference_M1_      440 KAPMKETKRQLPSTGETANPFFTAAALTVMATAGVAAVVKRKEEN    484
M5_sequence       195 -----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210916-185905-0776-44155798-plm.asequence
#   -bsequence emboss_needle-I20210916-185905-0776-44155798-plm.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
# Aligned_sequences: 2
# 1: Reference_M1_sequence
# 2: M8_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 490
# Identity: 124/490 (25.3%)
# Similarity: 165/490 (33.7%)
# Gaps: 258/490 (52.7%)
# Score: 449.5
=====

Reference_M1_      1 MAKNNNTNRHYSLRKLKTGTASAVALTVLGAGFANQTEVKANGD----G      45
| | : . | : . | | | | | | | | | | | | | : . | | | | | | | | | | | . .
M8_sequence       1 MARKDTNKQYSLRKLKTGTASAVAVAVLGA GAGFANQTEVKAESPKSHSIS 50

Reference_M1_      46 NPREVIEDLAANNPAIQNIRLYENKDLKARLENAMESVAGRDFKRAEELE 95
| . . : | . : | . | . : | | . | | . : | . . : | . . . | . : .
M8_sequence       51 NNEQLINEL---NDLIE-----ENNDLKDKIARNLDDL---LDNTREKD 87

Reference_M1_      96 KAKQALEDQRKDLETKLKE LQQDYDLAKESTS WDRQRLEKE EEEKKALE 145
. . . : | . . : | . . | . . | . . | . . | . . | : | . | : . | |
M8_sequence       88 PQYRALMGENDLREKEGKYQDKI-----KKLEEKEKNLE 122

Reference_M1_      146 LAIDQASRDYHRATALE -KELEEKKKALELAIDQASQDYNRANVLEKELE 194
. . . . : | . . : . . | . . : | . . | . . | . . | . . | . . | . . | . .
M8_sequence       123 KKSEDVERHYLKKLDQEHKEQQERQKNLE-----ELE 154

Reference_M1_      195 TITREQEINRNLLGNAKLELDQLSSEKEQLTIEKAKLEEKQISDASRQS 244
. . . . : | . . : . . | . . | . . | . . | . . | . . | . . | . . | . .
M8_sequence       155 R-----QSQREIDKRYQEQLQ---KQQQLETEKQISEASRK 188

Reference_M1_      245 LRRDL DASREAKKQVEKDLANLTAELDKVKEDKQISDASRQGLRRDIDAS 294
. . . . : | . . : . . | . . : | . . | . . | . . | . . | . . | . . | . .
M8_sequence       189 LSRDLEASRAAKDLE-----AEHQKLKEEKQISDASRQGLSRDLEAS 231

Reference_M1_      295 REAKKQVEKDLANLTAELDKVKEEKQISDASRQGLRRDIDASREAKKQVE 344
| . . . . : |
M8_sequence       232 REAKKKV----- 238

Reference_M1_      345 KALEEANSKLAALKLNKELEESKKLTKEKEKAELQAKLEAEAKALKEQLA 394
M8_sequence       239 ----- 238

Reference_M1_      395 KQAEELAKL RAGKASDSQTPDTKPGNKAVPGKGQAPQAGTKPNQNKA PMK 444
M8_sequence       239 ----- 238

Reference_M1_      445 ETKRQLPSTGETANPFFTAAALTVMATAGVAAVVKRKEEN 484
M8_sequence       239 ----- 238

=====
=====

```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210916-190019-0322-98296274-p2m.asequence
#   -bsequence emboss_needle-I20210916-190019-0322-98296274-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####
=====

# Aligned_sequences: 2
# 1: Reference_M1_sequence
# 2: M12_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 587
# Identity:      318/587 (54.2%)
# Similarity:    371/587 (63.2%)
# Gaps:          115/587 (19.6%)
# Score:         1326.5
=====

Reference_M1_      1 MAKNNNTNRHYSRLKLKTGTASAVAVALTVLGAGFANQTEVKANGDGNPREV        50
| || . | | | | | | | | | | | | | | | : | | | . . . . : | : .
M12_sequence      1 MAKNTTNRHYSRLKLKTGTASAVAVALTVVGAGLVAGQTVRADH-----       43

Reference_M1_      51 IEDLAANNPAIQNI----RLR-----YENKD-----        72
. | . | . . . . : | : | : | : | : | : | : | | : | | : | .
M12_sequence      44 -SDLVAEKQRLEDLGQKFERLQQRSELYLQQYYDNKSNGYKGDWYVQQLK        92

Reference_M1_      73 -LKARLENAM-EVAGRDFKRA-----EELEKAKQALEDQR        105
| . . . | . . | : | . . . | . . . | . . . | . . . | . . . | . . .
M12_sequence      93 MLNRDLEQAYNELSGEAHKDALGKLGIDNADLKA KIT ELEKSVEEKNDVL       142

Reference_M1_      106 KDLTAKLKEQQDYDLAKESTSWDRQRLKELEKK-----EALE        145
. . . . : | : | . . . . : | . . . : | . . : | . . . : | . . : | . .
M12_sequence      143 SQIKKELEEEAKD I QFGREVHAADLLRH KQEIAEKENVISKLNGELQPLK       192

Reference_M1_      146 LAIDQASR----DYHRATALEKEL----EEKKKALELAI---DQASQDYN        184
. . . : | . . | : . . : | : | : | . . | . . | . . | . . | : | :
M12_sequence      193 QKVDETDRNLQQEKKQKVLSLEQQLAVTKENAKKDFELAALGHQLADKEYN       242

Reference_M1_      185 RANVLEKELETITREQEINRNLLG--NAKLELDQLSSEKEQLTIEKAKLE        232
| . . . | . . . . : | . . | : . . | . . : | . . | . . | . . : | .
M12_sequence      243 -AKIALELESKLADAKKDFELAALGHQAHNEYQAKLAEKDG---QIKQLE       288

Reference_M1_      233 EEKQISDASRQLRRDLDASREAKKQVEKDLANLTAEELDKVKEDKQ----        278
| : | . | . | : . | . | : | : | . | : | : | . | . | . | . | . | . |
M12_sequence      289 EQKQILDASRKGTARDLEAVRQACKATEAELNNLKAELAKVTEQKQILDA       338

Reference_M1_      279 -----ISDASRQGLRRDLDASREA          297
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
M12_sequence      339 SRKGTARDLEAVRKAKAQVEAALKQLEEQN KISEASRQGLRRDLDASREA       388

Reference_M1_      298 KKQVEKDLANLTAEELDKVKEEKQISDASRQGLRRDLDASREAKKQVEKAL        347
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
M12_sequence      389 KKQVEKDLANLTAEELDKVKEEKQISDASRQGLRRDLDASREAKKQVEKAL       438

Reference_M1_      348 EEANSKLAALAEKL NKELEESKKLTEKEKAELQAKLEAEAKALKEQLAKQA        397
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
M12_sequence      439 EEANSKLAALAEKL NKELEESKKLTEKEKAELQAKLEAEAKALKEQLAKQA       488

Reference_M1_      398 EELAKL RAGKASD SQT P D T KPGNKAVPGKGQAPQAGT K P N Q N K A P M K E T K        447
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
M12_sequence      489 EELAKL RAGKASD SQT P D A KPGNKAVPGKGQAPQAGT K P N Q N K A P M K E T K       538

Reference_M1_      448 RQLPSTGETANPFFTAAALTVMATAGVAAVVKRKEEN        484
| : | : | : | . | : | : | : | : | : | : | : | : | : | : |
M12_sequence      539 RQLPSTGEAANPFFTAAALTVMATAGVAAVVKRKEEN        575
#-----
#-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210916-190109-0733-82553029-p2m.asequence
#   -bsequence emboss_needle-I20210916-190109-0733-82553029-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Reference_M1_sequence
# 2: M28_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 489
# Identity: 217/489 (44.4%)
# Similarity: 275/489 (56.2%)
# Gaps: 136/489 (27.8%)
# Score: 820.5
#
#=====

Reference_M1_      1 MAKNNNTNRHYSLRKLKLTGTASAVALTVLGAGFANQTEVKANGDGNPREV
                   ||::|:||:..|||||||||:|||||:|||||||:|
M28_sequence       1 MARKDTNKQYSLRKLKTGTASAVAVAVLGAQFANQTEVK-----
Reference_M1_      51 IEDLAANNPAIQNIRLRYENKDLKARLENAMESVAGRDFKRAEELEKAKQA
                   ||::| | ..|.....| .|...|....|.
M28_sequence       41 ----AAESP-----KSTETSGANGADK
Reference_M1_      101 LEDQRKDLETKLKELQQDY---DLAKESTS WDRQRLE-KELEEKKEALE
                   |.|....|.:::|:::| | ..|...|...|.| ::|..|:....|.
M28_sequence       58 LADAYNTLLTEHEKLRLDEYYTLIDAKEEPRYKALRGENQDLREKEGKYQ
Reference_M1_      146 LAIDQASRDYH RATALEKELEEKKKALELAIDQASQDYNRANVLEKELET
                   ..| | ..|:|:|..| | ..|:|..| | ..|:|:| |
M28_sequence       108 DKI-----KKLEEEKNLE---KKSEDVER-HYLKK---
Reference_M1_      196 ITREQEINRNLLGNAKLELDQLSSEKEQLTIEKAKLEEEEKQISDASRQL
                   |||...|:| | ..|:|:| ..|:|:| | ..|:|:| |
M28_sequence       135 -----LDQEHHKEQE-----ERQKNLEELERQS-
Reference_M1_      246 RRDL DASREAKKQVEKDLANLT AELDKVKEDKQISDASRQGLRDL DASR
                   :|::| | ..|:|:| ..| | ..|:|:| | ..|:|:| |
M28_sequence       157 QREID--KRYQEQLQKQ-----QOLETEKQISEASRKSLSRDLEASR
Reference_M1_      296 EAKKQVEKDLANLT AELDKVKEEKQISDASRQGLRDL DASREAKKQVEK
                   .|..|..| | ..|:|:| | ..|:|:| | ..|:|:| | ..|:|:| |
M28_sequence       197 AAKKDLE-----AEHQKLKEEKQISDASRQGLSRDLEASREAKKKVEA
Reference_M1_      346 ALEEANSKIAALEKLNKELEESKKL TEKEKAELQAKLEAEAKALKEQLAK
                   .|..|:| | ..|:|:| | ..|:|:| | ..|:|:| | ..|:|:| |
M28_sequence       240 DLAEANSKIQALEKLNKELEEGKKL SEKEKAELQARLEAEAKALKEQLAK
Reference_M1_      396 QAEELAKL RAGKASDSQT PDTD KPGNKA VPGKGQAPQAGTKPNQNQKAPMKE
                   |||..|:| | ..|:|:| | ..|:|:| | ..|:|:| | ..|:|:| |
M28_sequence       290 QAEELAKLKG----NQTPNAK-----VAPQA---NRSRSAMTQ
Reference_M1_      446 TKRQLPSTGETANPFFTAAALTVMATAGVA VVVKRKEEN      484
                   .|..|:| | ..|:|:| | ..|:|:| | ..|:|:| | ..|:|:| |
M28_sequence       321 QKRTLPSTGEAANPFFTAAAATVMVSAGMLA-LKRKEEN      358
#-----
#-----
```

```
#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210916-190152-0642-38607375-p2m.asequence
#   -bsequence emboss_needle-I20210916-190152-0642-38607375-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
# Aligned_sequences: 2
# 1: Reference_M1_sequence
# 2: M75_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 485
# Identity: 245/485 (50.5%)
# Similarity: 306/485 (63.1%)
# Gaps: 99/485 (20.4%)
# Score: 1021.0
=====

Reference_M1_      1 MAKNNNTNRHYSLRKLKTGTASAVALTVLGAGFANQTEVKANGDGNPREV      50
                   ||:::||:|||||||||:|||||:|||||:|||||..:....|
M75_sequence       1 MARKDTNKQYSLRKLKTGTASAVAVAVLGAGFANQTEVKAAEERTFTE-      49

Reference_M1_      51 IEDLAANNPAIQNIRLRYENKDLKARLENAMENVAGRDFKRAEELEKAKQA     100
                   ||..||    ||.:
M75_sequence       50 -----LPYE----ARYK-----      57

Reference_M1_      101 LEDQRKDLETKLKELQQDYDLAKESTSWDRQRLEKELEEKALELAIDQ     150
                   :| ..|.:|.:|.....:|
M75_sequence       58 -----AW-----KSENDELRENYRRTLDK      76

Reference_M1_      151 ASRDYHRATALEKELEEKKKALELA-IDQASQDYNRANVLEKELETITRE     199
                   .:.:.:.|||: .:||.||| :.:.:.|||: .:||.:.|||
M75_sequence       77 FNTEQGKTTTRLEE--QNKKLHSELASVTETLSVTEAD--DKKIKDLTDR     122

Reference_M1_      200 QEINRNLLGNAKLELDQLSSEKEQLTIEKAKLEEEKQISDASRQSLRRDL     249
                   .:|.:|||.|.:.::|:|||:|. ....|:|||:|||:|||..|||
M75_sequence       123 DKISSNLLGNAKDQINKLTTEKDTLAEKAKKLEEDKQISDASRKSLSRDL     172

Reference_M1_      250 DASREAKKQVEKDLANLTAEELDKVKEDKQISDASRQGLRRDLDASREAKK     299
                   :|||.|||:||.:.|||...:|||:|||:|||||:|||:|||:|||:|||:|||:|||
M75_sequence       173 EASRAAKKELEANHQKLETEHKKLKEEKQISDASRQGLSRDLEASREAKK     222

Reference_M1_      300 QVEKDLANLTAEELDKVKKEEKQISDASRQGLRRDLDASREAKKQVEKALEE     349
                   :|||.|||..|||:|||:|||||:|||:|||:|||:|||:|||:|||:|||:|||
M75_sequence       223 KVEADLAALTAEHQKLKEEKQISDASRQGLSRDLEASREAKKKVEADLAE     272

Reference_M1_      350 ANSKLAALEKLNLKELEESKKLTEKEKAELQAKLEAEAKALKEQLAKQAEE     399
                   |||||.|||||.|||:|||:|||||:|||:|||||:|||:|||:|||:|||:|||
M75_sequence       273 ANSKLQALEKLNLKELEEGKKLSEKEKAELQARLEAEAKALKEQLAKQAEE     322

Reference_M1_      400 LAKLRAGKASDSQT PDT KPGNKAVPGKGQAPQAGTKPNQNPKAPMKTKRQ     449
                   |||||. :|||.||| .||| | :|||:|||:|||:|||:|||:|||:|||
M75_sequence       323 LAKLKG----NQTPNAK-----VAPQA---NRSRSAMTQQKRT     353

Reference_M1_      450 LPSTGETANPFFTAAALTVMATAGVAAVVKRKEEN      484
                   |||||.|||||.|||..:|||. :|||:|||:|||:|||:|||:|||
M75_sequence       354 LPSTGEAANPFFTAAAATVMVSAGMLA-LKRKEEN      387

=====
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210916-190239-0819-2068156-p2m.asequence
#   -bsequence emboss_needle-I20210916-190239-0819-2068156-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
# Aligned_sequences: 2
# 1: Reference_M1_sequence
# 2: M79_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 486
# Identity: 125/486 (25.7%)
# Similarity: 157/486 (32.3%)
# Gaps: 266/486 (54.7%)
# Score: 461.0
=====

Reference_M1_      1 MAKNNNTNRHYSRKLKLTGTASAVALTVLGAGFANQTEVKANGDGNPREV      50
                   ||:::||:|||||||||:|||||:|||||.....!:
M79_sequence       1 MARKDTNKQYSRKLKLTGTASAVAVAVLGAGFANQTEVKAAEIKKPQ--  48

Reference_M1_      51 IEDLAANNPAIQNIRLRYENKDLKARLENAMESVAGRDFKRAEELEKAKQ-
                   .|.||.||...||.:||:||...||...||:||.||..|
M79_sequence       49 -ADSAWNWPKEYNALLK-ENEELKVEREKYLSY-----ADDKEKDPQY  89

Reference_M1_      100 -ALEDQRKDLETKLKELQDYLAKESTSWDRQRLEKEEEKAELELAI 148
                   ||...:|||...:...|...:|||...:|||..|..|:
M79_sequence       90 RALMGENQDLRKREGQYQDKI-----EELEKERKEKQERQE--- 125

Reference_M1_      149 DQASRDYHRATALEKEEKKKALELAIDQASQDYNRANVLEKELETITR 198
                   |..|.|. ...:||.:|||:| ..|:|:
M79_sequence       126 -QLERQYQ--IEADKHYZEQQQK-----K-----K 145

Reference_M1_      199 EQEINRNLLGNAKLELDQLSSEKEQLTIEKAKLEEEQKISDASRQSLRRD 248
                   .|:| ..|:||...||..:|||...||.|||:
M79_sequence       146 HQQ-----EQQQLEAEKQKLAKDKQISDASRQGLSRD 177

Reference_M1_      249 LDASREAKKQVEKDLANLTAELDKVKEDKQISDASRQGLRRDLDASREAK 298
                   |:|||||:||.|||.||||..|:|||||...|||..|||:
M79_sequence       178 LEASREAKKKVEADLAALTAEHQKLKDQKISDASRQGLSRDLEA---- 222

Reference_M1_      299 KQVEKDLANLTAELDKVKEEKQISDASRQGLRRDLDASREAKQVEKALE 348
                   223 -----
M79_sequence       223 ----- 222

Reference_M1_      349 EANSKLAALEKLNKELEESKKLTEKEKAELQAKLEAEAKALKEQLAKQAE 398
                   223 -----
M79_sequence       223 ----- 222

Reference_M1_      399 ELAKLRAGKASDSQTPDTKPGNKAVPGKGQAPQAGTKPNQNPKPMKETKR 448
                   223 -----
M79_sequence       223 ----- 222

Reference_M1_      449 QLPSTGETANPFFTAAALTVMATAGVAAVVKRKEEN      484
                   223 -----
M79_sequence       223 ----- 222

-----
-----

```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210916-190331-0721-53653255-p2m.asequence
#   -bsequence emboss_needle-I20210916-190331-0721-53653255-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
# Aligned_sequences: 2
# 1: Reference_M1_sequence
# 2: M81_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 484
# Identity: 278/484 (57.4%)
# Similarity: 316/484 (65.3%)
# Gaps: 88/484 (18.2%)
# Score: 1204.5
=====

Reference_M1_      1 MAKNNNTNRHYSLRKLKTGTASAVALTVLGAGFANQTEVKANGDGNPREV      50
                  .:.:.||:|||||:||||||:|||||:|||||:|||||.|.
M81_sequence      1 MVRKDTNKQYSLRKLKTGTASAVAVAVLGAGFANQTEVKAAAGS-----      44

Reference_M1_      51 IEDLAANNPAIQNIRLRYENKDLKARLENAMESVAGRDFKRAEELEKAKQA      100
                   ..|.|..|...|..|:||:       ||:|..|..:|||
M81_sequence      45 ---EENVPKQQYNALWEENEDLR-----GRERKYIAKLEK----      76

Reference_M1_      101 LEDQRKDLETKLKELQQDYDLAKESTSWDRQRLEKELEEKKEALELAIDQ      150
                   :|:|..
M81_sequence      77 -----EEIQNG-----      82

Reference_M1_      151 ASRDYHRATALEKELEKKALELAIDQASQDYNRANVLEKELETITREQ      200
                   |||.|..:|||..|... | |       |:|...:|.|:.
M81_sequence      83 -----ELNEKNRKLEADIADL-QD---VIEDNDQEIKRKD      113

Reference_M1_      201 EINRNLLGNAKLELDQLSSEKEQLTIEKAKLEEKQISDASRQLSRDLD      250
                   .:....|.:|.:..|::|:|..|....|..|:|..|:|..|:|..|:.
M81_sequence      114 RMYEAFLKQSKDQVKDLTAEKDTLAEKAKKLEEDKQISDASRKSLSRDLE      163

Reference_M1_      251 ASREAKKQVEKDLANLTAELDKVKEDKQISDASRQLRRDLDASREAKKQ      300
                   .|..|..|:|.....|..|:|:|||||:|||||:|||||:|..|:|:|||||:
M81_sequence      164 GSRAAKKELEAKHQKLETEHQKLKEDKQISDASRQGLSRDLEASREAKKK      213

Reference_M1_      301 VEKDLANLTAEVDKVEEKQISDASRQGLRRDLDASREAKKQVEKALEEA      350
                   |||..|..|..|:|:|||||:|||||:|||||:|..|:|:|..|:|..|:
M81_sequence      214 VEADLAALTAEHQKLKEEKQISDASRQGLSRDLEASREAKKKVEADLAEA      263

Reference_M1_      351 NSKLAALEKLNLKELEESKKLTEKEKAELQAKLEAEAKALKEQLAKQAEEL      400
                   |||||.|||||:|..|:|:|||||:|||||:|||||:|..|:|..|:|..|:
M81_sequence      264 NSKLQALEKLNLKELEEGKKLSEKEKAELQARLEAEAKALKEQLAKQAEEL      313

Reference_M1_      401 AKLRAKGASDSQTPTKPGNKAVPGKGQAPQAGTKPNQNPKPMKETKRQL      450
                   |||:|..|..|..|..|:|..|:|:|:|..|:|:|:|..|:|..|:|..|:|..|:
M81_sequence      314 AKLRAKGASDSQTPTAKPGNKVVPGKGQAPQAGTKPNQNPKPMKETKRQL      363

Reference_M1_      451 PSTGETANPFFTAAALTVMATAGVAAVVKRKEEN      484
                   |||||.|||||:|..|:|..|:|..|:|..|:|..|:|..|:|..|:
M81_sequence      364 PSTGEAANPFFTAAAATVMVSAGMLA-LKRKEEN      396

=====
#-----
#-----

```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210916-190422-0059-21208464-p2m.asequence
#   -bsequence emboss_needle-I20210916-190422-0059-21208464-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
# Aligned_sequences: 2
# 1: Reference_M1_sequence
# 2: M87_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 487
# Identity: 131/487 (26.9%)
# Similarity: 175/487 (35.9%)
# Gaps: 252/487 (51.7%)
# Score: 481.0
=====

Reference_M1_      1 MAKNNNTNRHYSRKLKLTGTASAVALTVLGAGFANQTEVKANGDGNPREV      50
                   ||:::||:|||||||||:|||||||. :|||| .
M87_sequence       1 MARKDTNKQYSRKLKLTGTASAVAVAVLGAGFANQTEVKAE---SPREV      47

Reference_M1_      51 IEDLAANNPAIQNIRLRYENKDLKARLENAMESVAGRDFKRAEELEKAKQA     100
                   ..:||: . .:||. .:||| .:||| .
M87_sequence       48 TNELAAS-----VWKKKVEEAKE-----KASKLEK---Q      73

Reference_M1_      101 LEDQRKD---LETKLKELOQDYDLAKESTSWDRQRLEKEEEKKEALELA    147
                   ||:::|| :|||:||| :||| .||:||| .
M87_sequence       74 LEEAQKDYSEIEGKLEQFWHDYD-----KLEK--ENKEYASQLG     110

Reference_M1_      148 IDQASRDYHRATALEKELEEKKKALELAIDQASQDYNRANVLEKELETIT    197
                   ..:..:....|.||.||:....:||. :||| .||:||| .
M87_sequence       111 KNQEEREKLELEYLRKSDEEYKEHQYRQEER-----QKNLEELE     152

Reference_M1_      198 REQEINRNLLGNNAKLELDQLSSEKEQLTIEKAKLEEEKQISDASRQLSLR    247
                   |.: .||:|...|.:| .:||| .|||:|||:||| .
M87_sequence       153 RQN-----KREIDKRYQEQLQ--KQQQLETEKQISEASRKSLSR     189

Reference_M1_      248 DLDASREAKKQVEKDLANLTAELDKVKEDKQISDASRQLRRDLDASREA    297
                   ||:|||.|||:| .||.||:|||:|||:||| .|||:||| .
M87_sequence       190 DLEASRAAKKELE-----AEHQKLKEEKQISDASRKSLSRLEASREA    232

Reference_M1_      298 KKQVEKDLANLTAELDKVKKEEKQISDASRQLRRDLDASREAKKQVEKAL    347
                   ||:|||.
M87_sequence       233 KKKVEA-----      238

Reference_M1_      348 EEANSKLAALEKLNKELEESKKLTEKEKAELQAKLEAEAKALKEQLAKQA    397
M87_sequence       239 -----      238

Reference_M1_      398 EELAKLРАГKASDSQTРDTKPGNKA VPГKGQAPQAGTKPNQNKPМKETK    447
M87_sequence       239 -----      238

Reference_M1_      448 RQLPSTGETANPFFTAALTVMATAGVAVVKRKEEN      484
M87_sequence       239 -----      238

-----
-----

```



```
#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210916-190546-0891-61986443-p2m.asequence
#   -bsequence emboss_needle-I20210916-190546-0891-61986443-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
# Aligned_sequences: 2
# 1: Reference_M1_sequence
# 2: M179_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 498
# Identity: 294/498 (59.0%)
# Similarity: 332/498 (66.7%)
# Gaps: 79/498 (15.9%)
# Score: 1218.5
=====

Reference_M1_ 1 MAKNNNTNRHYSLRKLKTGTASVAVALTVLGAGFA-NQTEVKANGDGNPRE 49
..:|||||:|||||:|||||:|||||.|||.|. .....
M179_sequence 1 MVRKDTRHYSLRKLKTGTASVAVALSVLGAGLAVNQTEVSAK--SVTRS 48

Reference_M1_ 50 VIEDLAANNPAIQNIRLRYENKDLKARLE-NAMEVAGRDFKRAEELEKAK 98
.:.....|| ..|.:...|....| |....|....
M179_sequence 49 TAQDPDKSRQAI---TEYEVENHKLTQEKN-----LTNRN 81

Reference_M1_ 99 QALEDQRKDLETKLKELQQDYDLAKESTS WDRQRLEKEEKK----- 141
.|.|.:...:|.:|...|:|...|....|.:.|||:|:|:|:
M179_sequence 82 QELTDENGELKTANEALRQRGDTLFN---QRVKLEKQVQEKEHNNKTLK 127

Reference_M1_ 142 -EALELAIDQASRDYHRATALEKELEEKKKALELAIDQASQDYNRA---N 187
|.|.|...:...| .|.|:|:|.:| |||...|....|.. .
M179_sequence 128 IE NGELKTENGD-----LTKKLD ETRQ--ELANKQ QESKENEKT LNE 167

Reference_M1_ 188 VLEKEL-ETITREQEINRNLLGNAKLELDQLSSEKEQLTIEKAKLEEKQ 236
:|...: .:|.:| |...:|...|...|:|...|.|||.:|:
M179_sequence 168 LLEKTVKDKIAKEQE-NKETIGTLKLLDETVKDK----IAKEQRSKQ 210

Reference_M1_ 237 ISDASRQSLRRDLDASREAKKQVEKDLANLTAELDKVKEDKQISDASRQG 286
|...|...|...|.:|...|:|...|:|...|:|...|:|...|:
M179_sequence 211 -----DFGALKQELAKKEEQN KISDASRQG 235

Reference_M1_ 287 LRRDL DASREAKKQVEKDLANLTAELDKVKEEKQISDASRQGLRRDLDAS 336
|...|...|...|:|...|:|...|:|...|:|...|:|...|:|...|:|...|:|...|:|...|:
M179_sequence 236 LRRDL N ASREAKKQVEKDLANLTAELDKVKEEKQVSDASRQGLRRDLDAS 285

Reference_M1_ 337 REAKKQVEKALEEANSKLAALEKLNKELEESKKLTEKEKAELQAKLEAEA 386
|...|...|...|:|...|:|...|:|...|:|...|:|...|:|...|:|...|:|...|:|...|:
M179_sequence 286 REAKKQVEKALEEANSKLAALEKLNKELEESKKLTEKEKAELQAKLEAEA 335

Reference_M1_ 387 KALKEQLAKQAEELAKL RAGKASDSQT P DTKPGNKAVPGKGQAPQAGTKP 436
|...|...|...|:|...|:|...|:|...|:|...|:|...|:|...|:|...|:|...|:|...|:
M179_sequence 336 KALKEKLAKQAEELAKL RAGKASDSQT P DAKPGNKAVPGKGQAPQAGTKP 385

Reference_M1_ 437 NQN KAPM KETKRQ LPSTGETANPFFTAA ALTV MATAGV AAVV KRKEEN 484
|...|...|...|:|...|:|...|:|...|:|...|:|...|:|...|:|...|:|...|:|...|:
M179_sequence 386 NQN KAPM KETKRQ LPSTGEA ANPFFTAA ALTV MATAGV AAVV KRKKEN 433

#-----
#-----
```

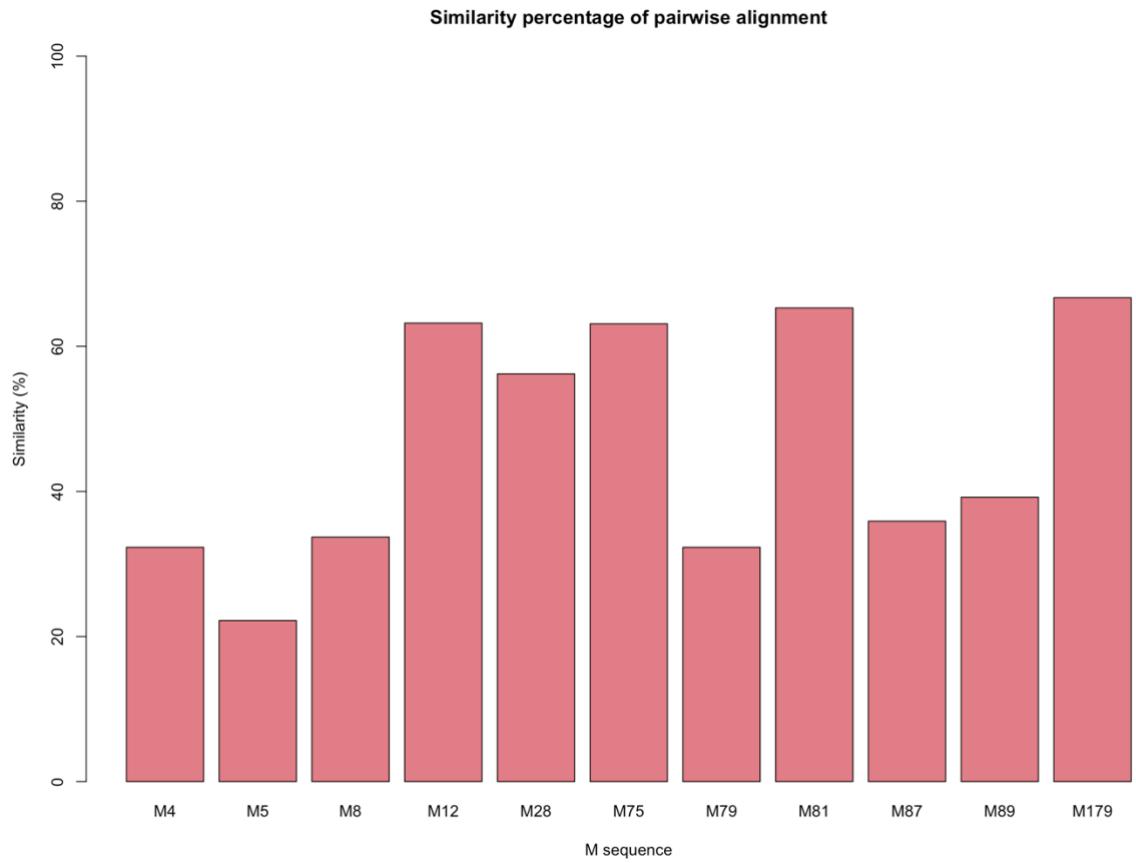


Figure S1. Similarity percentage for each M sequence aligned to M1 reference sequence

Binding peptides similarity in each sequence

Here, only the two peptides of M1 protein in the binding interface with ab25 *i.e.*, “KALELAIDQASQDYNRANVLEKE”, and “REAKKQVEKDLANLTAELDKVKE” were searched in each target M sequence to detect the similar patterns. Below, we report the result of the alignment and the similarity percentage with each of the detected peptides.

```
#####
# Program: needle
# Commandline: needle
# -auto
# -stdout
# -asequence emboss_needle-I20210917-134146-0008-34466034-p1m.asequence
# -bsequence emboss_needle-I20210917-134146-0008-34466034-p1m.bsequence
# -datafile EBLOSUM62
# -gapopen 10.0
# -gapextend 0.5
# -endopen 10.0
# -endextend 0.5
# -aformat3 pair
# -sprotein1
# -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep1
# 2: M4_similar_p1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 26
# Identity:      10/26 (38.5%)
# Similarity:    12/26 (46.2%)
# Gaps:          3/26 (11.5%)
# Score: 29.0
#
#
=====

Binding_pep1      1 KALELAIDQASQDYN--RANVLEKE      23
                  ||.|:...||...:|  ..|.|||
M4_similar_p1     1 KAAEIKKPQADSAWNWPKEYNALLKE      26
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-134735-0837-62507407-p1m.asequence
#   -bsequence emboss_needle-I20210917-134735-0837-62507407-p1m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep2
# 2: M4_similar_p2
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      17/23 (73.9%)
# Similarity:    19/23 (82.6%)
# Gaps:          0/23 ( 0.0%)
# Score:         76.0
#
#
=====

Binding_pep2      1 REAKKQVEKDLANLTAELDKVKE      23
                  |||||:||.|||.||||..|:||
M4_similar_p2     1 REAKKKVEADLAALTAEHQKLKE      23

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-135430-0963-60062530-p2m.asequence
#   -bsequence emboss_needle-I20210917-135430-0963-60062530-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep1
# 2: M5_similar_p1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      6/23 (26.1%)
# Similarity:    10/23 (43.5%)
# Gaps:          0/23 ( 0.0%)
# Score: 17.0
#
#
=====

Binding_pep1      1 KALELAIDQASQDYNRANVLEKE      23
                  |||...:::..:|.....|||
M5_similar_p1     1 KALNELLEKTVKDKIAKEQENKE      23

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-135546-0415-63161967-p2m.asequence
#   -bsequence emboss_needle-I20210917-135546-0415-63161967-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep2
# 2: M5_similar_p2
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      3/23 (13.0%)
# Similarity:    8/23 (34.8%)
# Gaps:          10/23 (43.5%)
# Score:         18.0
#
#
=====

Binding_pep2      1 REAKKQVEKDLANLTAELDKVKE      23
                  ...|.:.|:::|
M5_similar_p2     1 STQKETLEREVDQN-----      13

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-135644-0668-46832847-p2m.asequence
#   -bsequence emboss_needle-I20210917-135644-0668-46832847-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep1
# 2: M8_similar_p1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 25
# Identity:      6/25 (24.0%)
# Similarity:    13/25 (52.0%)
# Gaps:          2/25 ( 8.0%)
# Score: 17.5
#
#
=====

Binding_pep1      1 KALELAIDQASQDYNR--ANVLEKE      23
                  :..|...:::.|:|..|  ...|::|
M8_similar_p1     1 EEKEKNLEKKSEDVERHYLKKLDQE    25

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-135804-0655-27464049-p2m.asequence
#   -bsequence emboss_needle-I20210917-135804-0655-27464049-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep2
# 2: M8_similar_p2
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      10/23 (43.5%)
# Similarity:    13/23 (56.5%)
# Gaps:          3/23 (13.0%)
# Score:         34.0
#
#
=====

Binding_pep2      1 REAKKQVEKDLANLTAELDKVKE      23
                  .||.:..:||  ||.||..|:|| |
M8_similar_p2     1 LEASRAAKKD---LEAEHQKLKE      20
                  -----|-----|-----|-----|-----|
```

#-----
#-----

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-135919-0762-92494436-p2m.asequence
#   -bsequence emboss_needle-I20210917-135919-0762-92494436-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep1
# 2: M12_similar_p1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      7/23 (30.4%)
# Similarity:    12/23 (52.2%)
# Gaps:          0/23 ( 0.0%)
# Score: 29.0
#
#
=====

Binding_pep1      1 KALELAIDQASQDYNRANVLEKE      23
                  |.:|.|:::|:.....||.|||
M12_similar_p     1 KQVEKALEEANSKLAALKEKLNE      23

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-140032-0915-10411681-p2m.asequence
#   -bsequence emboss_needle-I20210917-140032-0915-10411681-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep2
# 2: M12_similar_p2
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      23/23 (100.0%)
# Similarity:    23/23 (100.0%)
# Gaps:          0/23 ( 0.0%)
# Score:         110.0
#
#
=====

Binding_pep2      1 REAKKQVEKDLANLTAELDKVKE      23
                  ||||||| ||||| ||||| ||||| |||||
M12_similar_p     1 REAKKQVEKDLANLTAELDKVKE      23

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-140131-0771-94163540-p2m.asequence
#   -bsequence emboss_needle-I20210917-140131-0771-94163540-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep1
# 2: M28_similar_p1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      6/23 (26.1%)
# Similarity:    10/23 (43.5%)
# Gaps:          0/23 ( 0.0%)
# Score: 22.0
#
#
=====

Binding_pep1      1 KALELAIDQASQDYNRANVLEKE      23
                  |.:|...:.|:.....|.|||
M28_similar_p     1 KKVEADLAEANSKLQALEKLNKE      23

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-140229-0375-95050436-p2m.asequence
#   -bsequence emboss_needle-I20210917-140229-0375-95050436-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep2
# 2: M28_similar_p2
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      11/23 (47.8%)
# Similarity:    17/23 (73.9%)
# Gaps:          0/23 ( 0.0%)
# Score:         54.0
#
#
=====

Binding_pep2      1 REAKKQVEKDLANLTAELDKVKE      23
                  |||||:||.|||...::|..:::
M28_similar_p     1 REAKKKVEADLAEANSKLQALEK      23

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-140324-0121-43219167-p2m.asequence
#   -bsequence emboss_needle-I20210917-140324-0121-43219167-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep1
# 2: M75_similar_p1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      6/23 (26.1%)
# Similarity:    10/23 (43.5%)
# Gaps:          0/23 ( 0.0%)
# Score: 22.0
#
#
=====

Binding_pep1      1 KALELAIDQASQDYNRANVLEKE      23
                  |.:|...:.|:.....|.|||
M75_similar_p     1 KKVEADLAEANSKLQALEKLNKE      23

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-140421-0869-93139689-p2m.asequence
#   -bsequence emboss_needle-I20210917-140421-0869-93139689-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep2
# 2: M75_similar_p2
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      17/23 (73.9%)
# Similarity:    19/23 (82.6%)
# Gaps:          0/23 ( 0.0%)
# Score:         76.0
#
#
=====

Binding_pep2      1 REAKKQVEKDLANLTAELDKVKE      23
                  |||||:||.|||.||||..|:||
M75_similar_p     1 REAKKKVEADLAALTAEHQKLKE      23

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-140533-0491-60880675-p1m.asequence
#   -bsequence emboss_needle-I20210917-140533-0491-60880675-p1m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep1
# 2: M79_similar_p1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 26
# Identity:      10/26 (38.5%)
# Similarity:    12/26 (46.2%)
# Gaps:          3/26 (11.5%)
# Score:         29.0
#
#
=====

Binding_pep1      1 KALELAIDQASQDYN--RANVLEKE      23
                  ||.|:...||...:|  ..|.||.|||
M79_similar_p     1 KAAEIKKPQADSAWNWPKEYNALLKE      26
                  -----
```

#-----
#-----

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-140643-0717-60803203-p2m.asequence
#   -bsequence emboss_needle-I20210917-140643-0717-60803203-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep2
# 2: M79_similar_p2
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      17/23 (73.9%)
# Similarity:    19/23 (82.6%)
# Gaps:          0/23 ( 0.0%)
# Score:         76.0
#
#
=====

Binding_pep2      1 REAKKQVEKDLANLTAELDKVKE      23
                  |||||:||.|||.||||..|:||
M79_similar_p     1 REAKKKVEADLAALTAEHQKLKE      23

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-140742-0256-53359928-p2m.asequence
#   -bsequence emboss_needle-I20210917-140742-0256-53359928-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep1
# 2: M81_similar_p1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      6/23 (26.1%)
# Similarity:    10/23 (43.5%)
# Gaps:          0/23 ( 0.0%)
# Score: 22.0
#
#
=====

Binding_pep1      1 KALELAIDQASQDYNRANVLEKE      23
                  |.:|...:.|:.....|.|||
M81_similar_p     1 KKVEADLAEANSKLQALEKLNKE      23

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-140900-0106-40553270-p2m.asequence
#   -bsequence emboss_needle-I20210917-140900-0106-40553270-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep2
# 2: M81_similar_p2
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      17/23 (73.9%)
# Similarity:    19/23 (82.6%)
# Gaps:          0/23 ( 0.0%)
# Score:         76.0
#
#
=====

Binding_pep2      1 REAKKQVEKDLANLTAELDKVKE      23
                  |||||:||.|||.||||..|:||
M81_similar_p     1 REAKKKVEADLAALTAEHQKLKE      23

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-140947-0333-77671278-p1m.asequence
#   -bsequence emboss_needle-I20210917-140947-0333-77671278-p1m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep1
# 2: M87_similar_p1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 34
# Identity:      9/34 (26.5%)
# Similarity:    15/34 (44.1%)
# Gaps:          11/34 (32.4%)
# Score:         32.0
#
#
=====

Binding_pep1      1 KALELAIDQASQDYNRA-----NVLEKE      23
                  ...||...::|.:||:..           :|||||
M87_similar_p     1 SKLEKQLEEAQKDYSIEGKLEQFWHDYDKLEKE      34

-----
-----

```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-141049-0262-1155708-p2m.asequence
#   -bsequence emboss_needle-I20210917-141049-0262-1155708-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep2
# 2: M87_similar_p2
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      10/23 (43.5%)
# Similarity:    13/23 (56.5%)
# Gaps:          7/23 (30.4%)
# Score: 35.0
#
#
=====

Binding_pep2      1 REAKKQVEKDLANLTAELDKVKE      23
                  ||||||:||       ||..|:||
M87_similar_p     1 RAAKKELE-----AEHQKLKE      16

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-141135-0412-59042471-p2m.asequence
#   -bsequence emboss_needle-I20210917-141135-0412-59042471-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep1
# 2: M89_similar_p1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      7/23 (30.4%)
# Similarity:    12/23 (52.2%)
# Gaps:          2/23 ( 8.7%)
# Score: 20.5
#
#
=====

Binding_pep1      1 KALELAIDQASQDYNRANVLEKE      23
                  ||.|.:...:.|..|::|
M89_similar_p     1 KAKEKSSENVERHYLRK--LDQE      21

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-141215-0123-81889031-p1m.asequence
#   -bsequence emboss_needle-I20210917-141215-0123-81889031-p1m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep2
# 2: M89_similar_p2
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      17/23 (73.9%)
# Similarity:    19/23 (82.6%)
# Gaps:          0/23 ( 0.0%)
# Score:         76.0
#
#
=====

Binding_pep2      1 REAKKQVEKDLANLTAELDKVKE      23
                  |||||:||.|||.||||..|:||
M89_similar_p     1 REAKKKVEADLAALTAEHQKLKE      23

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-141302-0615-69440915-p1m.asequence
#   -bsequence emboss_needle-I20210917-141302-0615-69440915-p1m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep1
# 2: M179_similar_p1
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      7/23 (30.4%)
# Similarity:    12/23 (52.2%)
# Gaps:          0/23 ( 0.0%)
# Score: 29.0
#
#
=====

Binding_pep1      1 KALELAIDQASQDYNRANVLEKE      23
                  |.:|.|:::|:.....|.|||
M179_similar_     1 KQVEKALEEANSKLAALKEKLNKE      23

-----
-----
```

```

#####
# Program: needle
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20210917-141404-0386-78941509-p1m.asequence
#   -bsequence emboss_needle-I20210917-141404-0386-78941509-p1m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Binding_pep2
# 2: M179_similar_p2
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 23
# Identity:      23/23 (100.0%)
# Similarity:    23/23 (100.0%)
# Gaps:          0/23 ( 0.0%)
# Score:         110.0
#
#
=====

Binding_pep2      1 REAKKQVEKDLANLTAELDKVKE      23
                  ||||||| ||||| ||||| ||||| |
M179_similar_     1 REAKKQVEKDLANLTAELDKVKE      23
                  ||||||| ||||| ||||| ||||| |

-----#
-----#

```

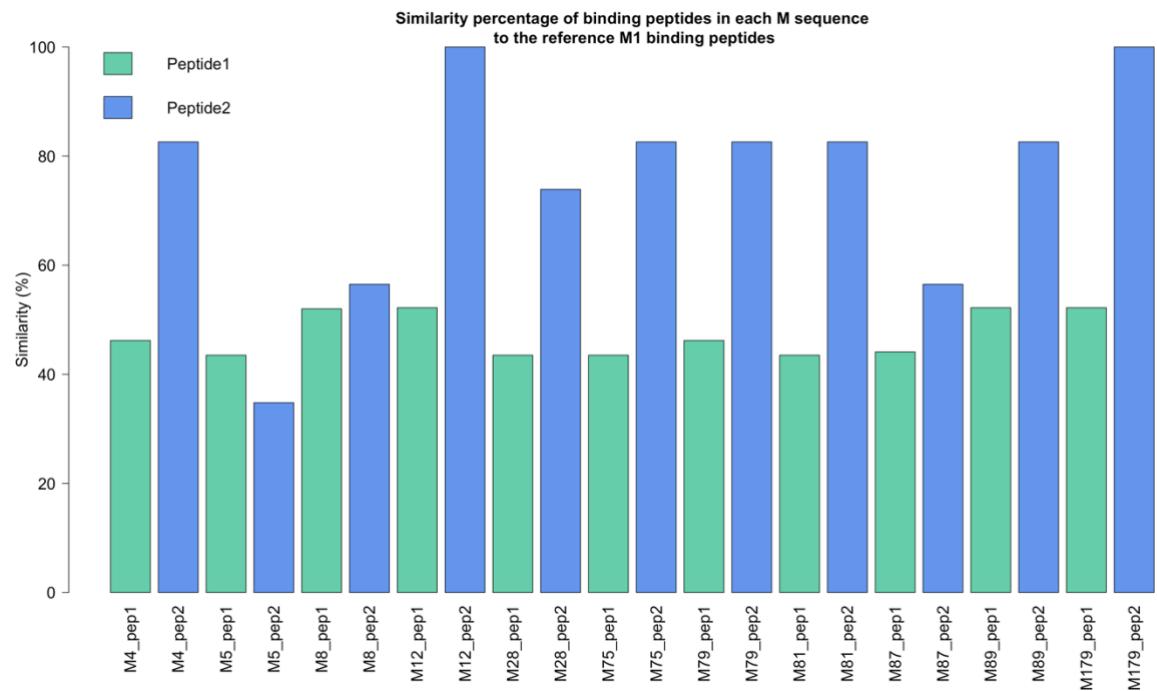


Figure S2. Similarity percentage for binding peptides of M1 against most similar peptide in each M sequence

Sequence similarity with M79 partial sequence as the reference

Here, the partial sequence of M79 as detected from sequencing data, is pairwise aligned against the rest of M proteins in this study. For those, M proteins that the full-length sequence is detected, we truncated the sequence to obtain a valid similarity score.

```
#####
# Program: needle
# Rundate: Wed  6 Oct 2021 09:49:10
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20211006-094725-0867-6737632-p1m.asequence
#   -bsequence emboss_needle-I20211006-094725-0867-6737632-p1m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Ref_M79_sequence
# 2: M1_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 295
# Identity:    125/295 (42.4%)
# Similarity: 157/295 (53.2%)
# Gaps:        75/295 (25.4%)
# Score:      461.0
#
# =====

Ref_M79_seque      1 MARKDTNKQYSLRKLKGTGTTASVAVAVAVLGAGFANQTEVKAAEIKKPQ--          48
                   ||:::||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|
M1_sequence       1 MAKNNNTNRHYSLRKLKGTGTTASVAVALTVLGAGFANQTEVKANGDGNPREV          50

Ref_M79_seque      49 -ADSAWNWPKEYNALLK-ENEELKVVEREKYLSY-----ADDKEKDPQY          89
                   .||.|.|...||..||:||:|||...|....|:||:..||..|.
M1_sequence       51 IEDLAANNPAIQNIRLRYENKDLKARLENAMESAGRDFKRAEELEKAKQ-          99

Ref_M79_seque      90 RALMGENQDLRKREGQQYQDKI-----EELEKERKEKQERQE---          125
                   ||....:||....:|...|..|.:|||..:||:|..|.
M1_sequence       100 -ALEDQRKDLETKLKELQQDYDLAKEESTSWDRQRLEKELEKKAEALELAI         148

Ref_M79_seque     126 -QLERQYQ--IEADKHQYQEQQK-----K          145
                   |..|..|...:|..:|..:|..|.
M1_sequence       149 DQASRDYHRATALEKELEEKKKALELAIDQASQDYNRANVLEKELETITR         198

Ref_M79_seque     146 HQQ-----EQQQLEAEKQKLAKDKQISDASRQGLSRD          177
                   .|:|..|.:||..|..|..|..|:|||..|..|..|..|.
M1_sequence       199 EQEINRNLLGNAKLELDQLSSEKEQLTIEKAKLEEKQISDASRQSLRRD         248

Ref_M79_seque     178 LEASREAKKKVEADLAALTAEHQKLKEDKQISDASRQGLSRDLEA          222
                   |:|||||||:||..|||..|||..|:|||..|..|..|..|:|
M1_sequence       249 LDASREAKKQVEKDLANLTAELDKVVEDKQISDASRQGLRRDLD          293

-----
-----
```

```
#####
# Program: needle
# Rundate: Wed  6 Oct 2021 09:51:02
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20211006-095058-0469-54470083-p2m.asequence
#   -bsequence emboss_needle-I20211006-095058-0469-54470083-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####
=====
```

```
#
# Aligned_sequences: 2
# 1: Ref_M79_sequence
# 2: M4_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 222
# Identity: 222/222 (100.0%)
# Similarity: 222/222 (100.0%)
# Gaps: 0/222 ( 0.0%)
# Score: 1113.0
#
#=====
```

Ref_M79_seque	1	MARKDTNKQYSRKLKTGTASVAVAVAVLGAGFANQTEVKAAEIKKPQAD	50
M4_sequence	1	MARKDTNKQYSRKLKTGTASVAVAVAVLGAGFANQTEVKAAEIKKPQAD	50
Ref_M79_seque	51	SAWNWPKEYNALLKENEELKVEREKYLSYADDKEKDPQYRALMGENQDLR	100
M4_sequence	51	SAWNWPKEYNALLKENEELKVEREKYLSYADDKEKDPQYRALMGENQDLR	100
Ref_M79_seque	101	KREGQYQDKIEELEKERKEKQERQEQLERQYQIEADKHYQEQQKKHQEQ	150
M4_sequence	101	KREGQYQDKIEELEKERKEKQERQEQLERQYQIEADKHYQEQQKKHQEQ	150
Ref_M79_seque	151	QQLEAEKQKLAKDKQISDASRQGLSRDLEASREAKKKVEADLAALTAEHQ	200
M4_sequence	151	QQLEAEKQKLAKDKQISDASRQGLSRDLEASREAKKKVEADLAALTAEHQ	200
Ref_M79_seque	201	KLKEDKQISDASRQGLSRDLEA 222	
M4_sequence	201	KLKEDKQISDASRQGLSRDLEA 222	

```
-----
-----
```

```
#####
# Program: needle
# Rundate: Wed  6 Oct 2021 09:44:41
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20211006-095144-0370-41504014-p1m.asequence
#   -bsequence emboss_needle-I20211006-095144-0370-41504014-p1m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####
=====
```

```
#
# Aligned_sequences: 2
# 1: Ref_M79_sequence
# 2: M5_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 231
# Identity:      70/231 (30.3%)
# Similarity:    112/231 (48.5%)
# Gaps:          46/231 (19.9%)
# Score:         219.5
#
#
```

```
=====
Ref_M79_seque      1 MARKDTNKQYSLRKLKTGTASVAVAVAVLGAGF-ANQTEVKAA---EIK      45
                   |||::|||..|||||||.|||||||:|||||. .|||.|.|| .|.
M5_sequence        1 MARENTNKHYSLRKLKKGTAASVAVSLVGAGLVVNTNEVSAAVTRGTIN      50
Ref_M79_seque      46 KPQ--ADSAWNWPKEYNALLKENEEELKVEREKYLSYADDKEKDQYRALM      93
                   .|| ..:.:.:.:.:.:.:.:.||.||.||.:. :.:.||.
M5_sequence        51 DPQRRAKEALDKYELENHDLKTKNEGLKTENEGLKT-----ENEGLK      91
Ref_M79_seque      94 GENQDLRKREGQQYQDKIEELEKERKEKQERQEQLERQYQIEADKHYQEQQ-
                   .||.:.||.:.:.:.:.:.:.||.||.||.:.||.:.||.:.||.:
M5_sequence        92 TENEGLKTEKKEHEAENDKLKQQRTLSTQKETLEREVQ---NTQYNNET      138
Ref_M79_seque      143 -QKKHQQEQQQLEAEKQKLAKDKQISDASRQGLSRDLEASREAKKKVEAD      191
                   .:.||.:.:.||.:.||.:.||.:.||.:.||.:.||.:.||.:.||.:.||.:
M5_sequence        139 LKIKNGDLTKELNKTRQELANKQQESKENEKALNELLE--KTVKDKI---      183
Ref_M79_seque      192 LAALTAEHQKLKEDKQISDASRQGLSRDLEA      222
                   |.:.||.||.|.
M5_sequence        184 -----AKEQENKETIG-----      194
```

```
=====
=====
```

```

#####
# Program: needle
# Rundate: Wed  6 Oct 2021 09:47:09
# Commandline: needle
#       -auto
#       -stdout
#       -asequence emboss_needle-I20211006-095238-0087-7711006-p1m.asequence
#       -bsequence emboss_needle-I20211006-095238-0087-7711006-p1m.bsequence
#       -datafile EBLOSUM62
#       -gapopen 10.0
#       -gapextend 0.5
#       -endopen 10.0
#       -endextend 0.5
#       -aformat3 pair
#       -sprotein1
#       -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Ref_M79_sequence
# 2: M8_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 256
# Identity:      149/256 (58.2%)
# Similarity:    170/256 (66.4%)
# Gaps:           52/256 (20.3%)
# Score:          641.0
#
#
=====

Ref_M79_seque 1 MARKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTEVKAAEIKKPQAD      50
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | . . | . . .
M8_sequence   1 MARKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTEVKAESPKSHSIS      50

Ref_M79_seque 51 SAWNWPKEYNALLKENEELKVEREKYLYSADD-KEKDPQYRALMGENQDL      99
: : . . . . . . : : . : . : . : . . . . . . . : . : : . . . . . . . .
M8_sequence   51 NNEQLINELNDLIEENNNDLKDKLARNLDLLDNTREKDPQYRALMGENQDL     100

Ref_M79_seque 100 RKREGQYQDKIEELE-----KERKEKQERQ---      124
| : | : | | | | : : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
M8_sequence   101 REKEGKYQDKIKKLEEKEKNLEKKSEDVERHYLKKLDQEHKEQQERQKNL     150

Ref_M79_seque 125 EQLERQYQIEADKHYQEQQKKHQQEQQQLEAEKQQLAKDKQISDASRQGL      174
| : | | | . . . | | | | | | . | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
M8_sequence   151 EELERQSQRIDEKRYSEQ---LQKQQQLETE-----KQISEASRKSL     189

Ref_M79_seque 175 SRDLEASREAKKKVEADLAALTAEHQKLKEDKQISDASRQGLSRDLEA--      222
| | | | | | . | | | : | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
M8_sequence   190 SRDLEASRAAKKDLE-----AEHQKLKEEKQISDASRQGLSRDLEASR     232

Ref_M79_seque 223 -----      222
M8_sequence   233 EAKKKV      238

=====
=====

```

```
#####
# Program: needle
# Rundate: Wed  6 Oct 2021 09:53:19
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20211006-095514-0934-92394355-p2m.asequence
#   -bsequence emboss_needle-I20211006-095514-0934-92394355-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####
=====
```

```
#
# Aligned_sequences: 2
# 1: Ref_M79_sequence
# 2: M12_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 323
# Identity:      89/323 (27.6%)
# Similarity:    134/323 (41.5%)
# Gaps:          117/323 (36.2%)
# Score:         270.5
#
#
```

Ref_M79_seque	1 MARKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTEVKAAEIKKPQAD	50
M12_sequence	1 MAKNTTNRHYSRKLKTGTASVAVALTVVGAGLVAGQTVR-----	41
Ref_M79_seque	51 SAWNWPKEYNALLKEN-----EELKVEREKYL-SYADDK-----	83
M12_sequence	42 -----DHSDLVAEKQRLEDLGQKFERLKQRSELYLQQYYDNKSNGYKG	84
Ref_M79_seque	84 -----EKDPQYRALMGE-----NQDLRKREGQYQDK	109
M12_sequence	85 DWYVQQLKMLNRDLEQAYNELSGEAHKDALGKLGINADLKAKITELEKS	134
Ref_M79_seque	110 IEE----LEKERKEKQE-----RQEQLERQYQI---	133
M12_sequence	135 VEEKNDVLSQIKKELEAEKDQFREVHAADLLRHKQEIAEKENVISKL	184
Ref_M79_seque	134 -----EADKHYQEQQKKHQQEQQQLEAEKQKLAKDKQISDASR	171
M12_sequence	185 NGELQPLKQKVDETDRNLQQEKQKVLSLEQQILAVTKENAKKDFELAALGH	234
Ref_M79_seque	172 QGLSRDLEA-----SREAKKKVEADLAALTAEH-----	199
M12_sequence	235 QLADKEYNAKIAELESKLADAKKDFELAALGHQAHNEYQAKLAEKDQI	284
Ref_M79_seque	200 QKLKEDKQISDASRQGLSRDLEA	222
M12_sequence	285 KQLEEQQKILDASRKGTARDLEA	307

```
=====
=====
```

```
#####
# Program: needle
# Rundate: Wed  6 Oct 2021 09:53:38
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20211006-100101-0387-61592505-p1m.asequence
#   -bsequence emboss_needle-I20211006-100101-0387-61592505-p1m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
```

```
# Aligned_sequences: 2
# 1: Ref_M79_sequence
# 2: M28_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 250
# Identity: 150/250 (60.0%)
# Similarity: 173/250 (69.2%)
# Gaps: 49/250 (19.6%)
# Score: 642.0
#
#
```

Ref_M79_seque	1 MARKDTNKQYSRKLKTGTASVAVAVAVLGAGFANQTEVKAAEIKKPQAD	50
M28_sequence	1 MARKDTNKQYSRKLKTGTASVAVAVAVLGAGFANQTEVKAAESPKSTET	50
Ref_M79_seque	51 SAWNWPK---EYNALLKENEELKVEREKYLSYADDKEKDQPQYRALMGENQ	97
M28_sequence	51 SANGADKLADAYNTLLTEHEKL---RDEYYTLIDAKEEEPRYKALRGENQ	97
Ref_M79_seque	98 DLRKREGQQYQDKIEELE-----KERKEKQERQ-	124
M28_sequence	98 DLREKEGKYQDKIKKLEEKEKNLEKKSEDVERHYLKKLDQEHKEQEERQK	147
Ref_M79_seque	125 --EQLERQYQIEADKHYQEQQKKHQEQQQLEAEKQKLAKDKQISDASRQ	172
M28_sequence	148 NLEELERQSREIDKRYQE---LQKQQQLETE-----KQISEASRK	186
Ref_M79_seque	173 GLSRDLEASREAKKKVEADLAALTAEHQKLKEDKQISDASRQGLSRDLEA	222
M28_sequence	187 SLSRDLEASRAAKKDLE-----AEHQKLKEEKQISDASRQGLSRDLEA	229

```
=====
=====
```

```
#####
# Program: needle
# Rundate: Wed  6 Oct 2021 09:58:18
# Commandline: needle
#      -auto
#      -stdout
#      -asequence emboss_needle-I20211006-100349-0215-55773762-p1m.asequence
#      -bsequence emboss_needle-I20211006-100349-0215-55773762-p1m.bsequence
#      -datafile EBLOSUM62
#      -gapopen 10.0
#      -gapextend 0.5
#      -endopen 10.0
#      -endextend 0.5
#      -aformat3 pair
#      -sprotein1
#      -sprotein2
# Align_format: pair
# Report_file: stdout
#####
```

```
=====
#
```

```
# Aligned_sequences: 2
# 1: Ref_M79_sequence
# 2: M75_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 224
# Identity: 112/224 (50.0%)
# Similarity: 143/224 (63.8%)
# Gaps: 10/224 ( 4.5%)
# Score: 450.0
#
#
```

```
=====
#
```

Ref_M79_seque	1 MARKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTEVKAAEIKKPQAD	50
	
M75_sequence	1 MARKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTEVKAAE---ERTF	47

Ref_M79_seque	51 SAWNWPKEYNALLKENEELKVEREKYLSYADDKEKDPQYRALMGENQDLR	100
	:.....:... : :....:.... . : :	
M75_sequence	48 TELPYEARKYAWKSENDELRENYRRTL---DKEFNTEQGKTTLEEQN-K	92

Ref_M79_seque	101 KREGQYQDKIEELEKERKEKQERQEQLERQYQIEAD--KHYQEQQKKHQQ	148
:....:....:....	
M75_sequence	93 KLHSELASVTETLTSVTEADDKKIKDLTDRDKISSNLLGNAKDQINKLTT	142

Ref_M79_seque	149 EQQQLEAEKQKLAKDKQISDASRQGLSRDLEASREAKKKVEADLAALTAE	198
	
M75_sequence	143 EKDTLAEKAKKLEEDKQISDASRKSLSRDLEASRAAKKELEANHQKLETE	192

Ref_M79_seque	199 HQKLKEDKQISDASRQGLSRDLEA	222
	: :	
M75_sequence	193 HKKLKEEKQISDASRQGLSRDLEA	216

```
=====
-----
```

```
#####
# Program: needle
# Rundate: Wed  6 Oct 2021 10:07:10
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20211006-100706-0795-49043706-p2m.asequence
#   -bsequence emboss_needle-I20211006-100706-0795-49043706-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
```

```
#
# Aligned_sequences: 2
# 1: Ref_M79_sequence
# 2: M81_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 223
# Identity:      124/223 (55.6%)
# Similarity:    148/223 (66.4%)
# Gaps:           18/223 ( 8.1%)
# Score:          509.5
#
#
```

```
=====
Ref_M79_seque      1 MARKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTEVKAЕIKKPQAD      50
                   |.|||||||||||||||||||||||||||||||||.|.
M81_sequence       1 MVRKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTEVKA-----AG      43
Ref_M79_seque      51 SAWNWPK-EYNALLKENEELKVEREKYLSYADDKEKDPQYRALMGENQDL     99
                   |..|.|| :||||.:|||:|:.....||:...:|| .|....|.:|:|
M81_sequence        44 SEENVPKQQYNALWEENEDILRGGRERKYIAKLEKEE--IQNGELNEKNRKL     91
Ref_M79_seque      100 RKREGQYQDKIEELEKERKEKQERQEQLERQYQIEADKHYQEQQKKHQQE    149
                   .....||.:||:||.|| :||.||:...|..:||.||....|
M81_sequence        92 EADIADLQDVIEDNDQEIKRK-----DRMYE AFL-KQSKDQVKDLTAE    133
Ref_M79_seque      150 QQQLEAEKQKLAKDKQISDASRQGLSRDLEASREAKKKVEADLAALTAEH    199
                   :..|...:||.||:|||||||:|||||||.||.|||:||.||...|..|||
M81_sequence        134 KDTLAEKAKKLEEDKQISDASRKSLSRDLEGSRAAKKELEAKHQKLETEH    183
Ref_M79_seque      200 QKLKEDKQISDASRQGLSRDLEA      222
                   |||||||||||||||||||||
M81_sequence        184 QKLKEDKQISDASRQGLSRDLEA      206
```

```
=====
=====
```

```
#####
# Program: needle
# Rundate: Wed  6 Oct 2021 10:10:16
# Commandline: needle
#      -auto
#      -stdout
#      -asequence emboss_needle-I20211006-101012-0859-31002285-p2m.asequence
#      -bsequence emboss_needle-I20211006-101012-0859-31002285-p2m.bsequence
#      -datafile EBLOSUM62
#      -gapopen 10.0
#      -gapextend 0.5
#      -endopen 10.0
#      -endextend 0.5
#      -aformat3 pair
#      -sprotein1
#      -sprotein2
# Align_format: pair
# Report_file: stdout
#####
```

```
=====
#
```

```
# Aligned_sequences: 2
# 1: Ref_M79_sequence
# 2: M87_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 246
# Identity: 131/246 (53.3%)
# Similarity: 159/246 (64.6%)
# Gaps: 42/246 (17.1%)
# Score: 516.5
#
#
```

```
=====
Ref_M79_seque   1 MARKDTNKQYSRKLKTGTASVAVAVAVLGAGFANQTEVK---AEIKKP 47
                  ||||||| ||||| ||||| | | | | | | | | | | | | | | | | .|:...
M87_sequence    1 MARKDTNKQYSRKLKTGTASVAVAVAVLGAGFANQTEVKAESPREVTNE 50
Ref_M79_seque   48 QADSAW-----NWPKEYNALLKENEELKVEREKYLSYADDKEK 85
                  .|.|. | .|...|:....|:..|:...|:....|..| |
M87_sequence    51 LAASVWKKKVVEEAKEKASKLEKQLEEAQKDYSEIEGKLEQFWHDYDKLEK 100
Ref_M79_seque   86 D-PQYRALMGENQDLR-KREGQQYQDKIEELEKE---RKEKQERQ---EQ 126
                  : ..|.:|:||:|..| |..|:|...|.|| |:|:||| | |:
M87_sequence    101 ENKEYASQLGKNQEEREKLELEYLRKSDEEYKEHQYRQEERQKNLEE 150
Ref_M79_seque   127 LERQYQIEADKHYQEQQKKHQQEQQLEAEKQKLAKDKQISDASRQGLSR 176
                  |||||.:..|..| | | | | |:| | | | | | | | | | | | | |
M87_sequence    151 LERQNKRKDQYQEQ----LQKQQQLETE-----KQISEASRKSLSR 189
Ref_M79_seque   177 DLEASREAKKKVEADLAALTAEHQKLKEDKQISDASRQGLSRDLEA 222
                  ||||||| .| | | :| | | | | | | | | | | | | | | | | |
M87_sequence    190 DLEASRAAKKELE-----AEHQKLKEEKQISDASRKSLSRDLEA 228
```

```
-----
-----
```

```
#####
# Program: needle
# Rundate: Wed  6 Oct 2021 10:12:06
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20211006-101051-0273-48298929-p1m.asequence
#   -bsequence emboss_needle-I20211006-101051-0273-48298929-p1m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
```

```
#
# Aligned_sequences: 2
# 1: Ref_M79_sequence
# 2: M89_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 222
# Identity:      91/222 (41.0%)
# Similarity:    124/222 (55.9%)
# Gaps:          61/222 (27.5%)
# Score:         364.0
#
#
```

```
=====
Ref_M79_seque     1 MARKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTEVKAAEIKKPQAD    50
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | . | |
M89_sequence      1 MARKDTNKQYSLRKLKTGTASVAVAVAVLGAGFANQTTVKA-----DSD    44
Ref_M79_seque     51 SAWNWPKEYNALLKENEELKVEREKYLSYADDKEKDPQYRALMGENQDLR   100
..           | . : | : | . | | . | : | . | | . | : | . | | . | . | . | . |
M89_sequence      45 NI-----NRSVSVK-----DNEKE---LHNKIADLE    67
Ref_M79_seque     101 KREGQYQDKIEELEKERKEKQERQEQLERQYQIEADKHYZEQQQKKHQQEQ  150
: . | : . | | : | : | . | ; : | : | . | : | . | : | | : | : | : | . | :
M89_sequence       68 EERGEHLDKIDELKEELKAKEKSSENVERHYLRKLDQEYKEQQER-QKNL  116
Ref_M79_seque     151 QQLEAEKQQLAKDKQISDASRQGLSRDLEASREAKKKVEADLAALTAEHQ  200
: : | | . : |           | | . : | : : . | . : | . | . | . | . | . | . | . |
M89_sequence      117 EELERQSQ-----REVEKRYQEQL---QKQQ    139
Ref_M79_seque     201 KLKEDKQISDASRQGLSRDLEA      222
: : . | | | : | | : | . | | | | | | | | | | | | | | | | | | | | | | | | |
M89_sequence      140 QLETEKQISEASRKSLSRDLEA      161
```

```
=====
=====
```

```

#####
# Program: needle
# Rundate: Wed  6 Oct 2021 10:13:16
# Commandline: needle
#   -auto
#   -stdout
#   -asequence emboss_needle-I20211006-101315-0196-46816129-p2m.asequence
#   -bsequence emboss_needle-I20211006-101315-0196-46816129-p2m.bsequence
#   -datafile EBLOSUM62
#   -gapopen 10.0
#   -gapextend 0.5
#   -endopen 10.0
#   -endextend 0.5
#   -aformat3 pair
#   -sprotein1
#   -sprotein2
# Align_format: pair
# Report_file: stdout
#####

=====
#
# Aligned_sequences: 2
# 1: Ref_M79_sequence
# 2: M179_sequence
# Matrix: EBLOSUM62
# Gap_penalty: 10.0
# Extend_penalty: 0.5
#
# Length: 289
# Identity:      118/289 (40.8%)
# Similarity:    154/289 (53.3%)
# Gaps:          72/289 (24.9%)
# Score:         447.0
#
#
=====

Ref_M79_seque      1 MARKDTNKQYSLRKLKTGTASVAVAVAVLGAGFA-NQTEVKAAE I-----        44
                   | .|||||:|.||||||||||||||:|||||.|||.|.|||.|.|.
M179_sequence     1 MVRKDTNRHYSLRKLKTGTASVAVALSVLGAGLAVNQTEVSAKS VTRSTA        50
                   | ..|..:..:...:|.|..|..|:|..|.
Ref_M79_seque     45 ---KKPQADSAW---NWPKEYNALLKENEEL-----        69
                   | ..|..:..:...:|.|..|..|:|..|.
M179_sequence     51 QDPDKSRQAITEYEVENHKLTQEKNALTNRNQELTDENGELKTANEALRQ        100
                   | ..|..:..:...:|.|..|..|:|..|.
Ref_M79_seque     70 -----KVEREKYLSYADDKEKDQPYRALMGENQDLRKREGQYQDKIE        111
                   :|.:|..:..:|.:..:..|..|..|:|..|..|:|
M179_sequence     101 RGDTLFNQRVKLEKQV---QEKEHNNKTLKIENGELKTENGDLTKLLD        145
                   | ..|..:..:|..:..|..:..|..|..|:|..|.
Ref_M79_seque     112 ELEKERKEKQERQEQLERQY---QIEADKHYQEQQK-----        144
                   | ..|..:..:|..:..|..:..|..|..|:|..|.
M179_sequence     146 ETRQELANKQQESKENEKTLNELLEKTVKDKIAKEQENKETIGTLKKLD        195
                   | ..|..:..:|..:..|..:..|..|..|:|..|.
Ref_M79_seque     145 -----KHQQEQQQLEAEKQQLAKDKQ---ISDASRQGLSRDLEASRE        183
                   | ..|..:..:|..:..|..:..|..|..|:|..|..|.
M179_sequence     196 ETVKDKIAKEQKSKQDFGALKQELAKKEEQN KISDASRQGLRRDLNASRE        245
                   | ..|..:..:|..:..|..:..|..|..|:|..|..|.
Ref_M79_seque     184 AKKKVEADLAALTAEHQKLKEDKQISDASRQGLSRDLEA        222
                   | ..|..:..:|..:..|..:..|..|..|:|..|..|.
M179_sequence     246 AKKQVEKDLANLTAELDKVKEEKQVSDASRQGLRRDLDA        284
                   | ..|..:..:|..:..|..:..|..|..|:|..|..|.

=====
=====

```

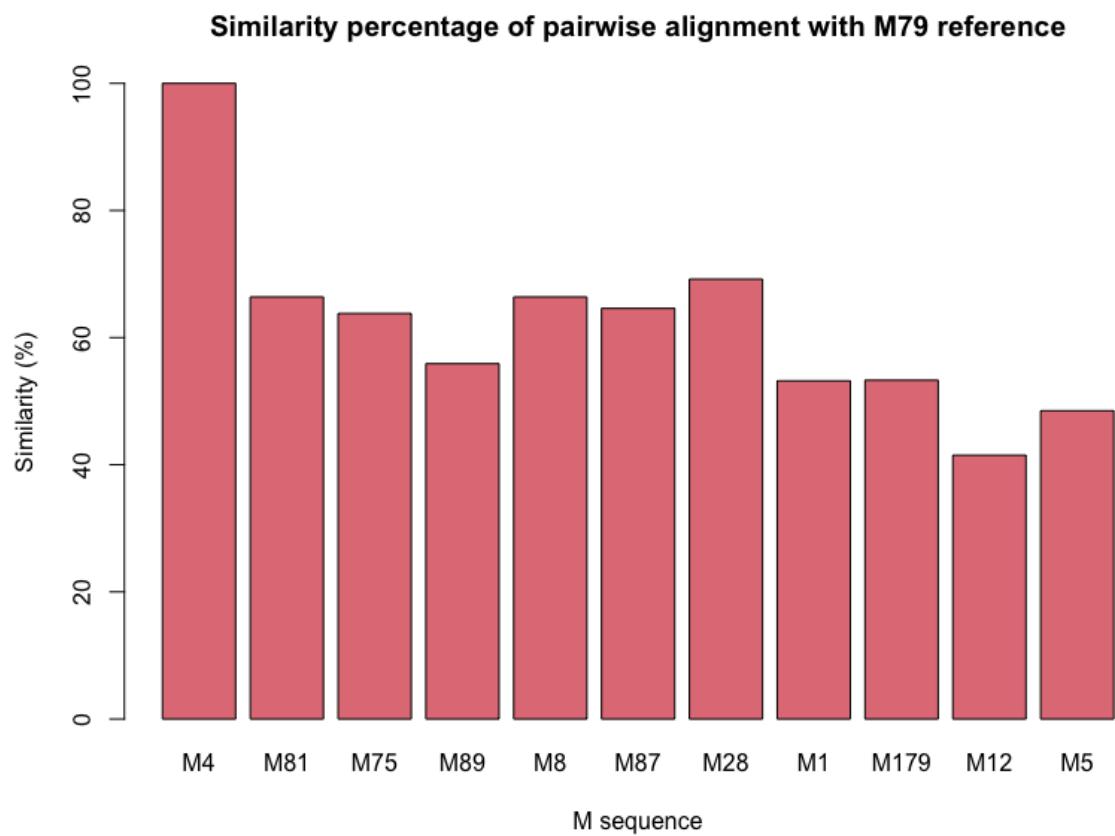


Figure S3. Similarity percentage for truncated M sequences against M79 partial reference sequence