

Elongation factor Tu is a multifunctional and processed moonlighting protein

Michael Widjaja^{1,†}, Kate Louise Harvey^{1,†}, Lisa Hagemann^{2,†}, Iain James Berry¹, Veronica Maria Jarocki¹, Benjamin Bernard Armando Raymond¹, Jessica Leigh Tacchi¹, Anna Gründel², Joel Ricky Steele¹, Matthew Paul Padula³, Ian George Charles⁴, Roger Dumke^{2,‡}, Steven Philip Djordjevic^{1,3,‡*}

¹ The itthree institute, University of Technology Sydney, PO Box 123, Broadway, NSW, 2007, Australia.

² Technische Universität Dresden, Medizinische Fakultät Carl Gustav Carus, Institut für Medizinische Mikrobiologie und Hygiene, Fetscherstrasse 74, 01307 Dresden, Germany.

³ Proteomics Core Facility, University of Technology Sydney, PO Box 123, Broadway, NSW, 2007, Australia.

⁴ Institute for Food Research, Norwich Research Park, Norwich, NR4 7UA, UK.

† These authors contributed equally to this work

‡ These contributors share senior authorship

*Corresponding Author:

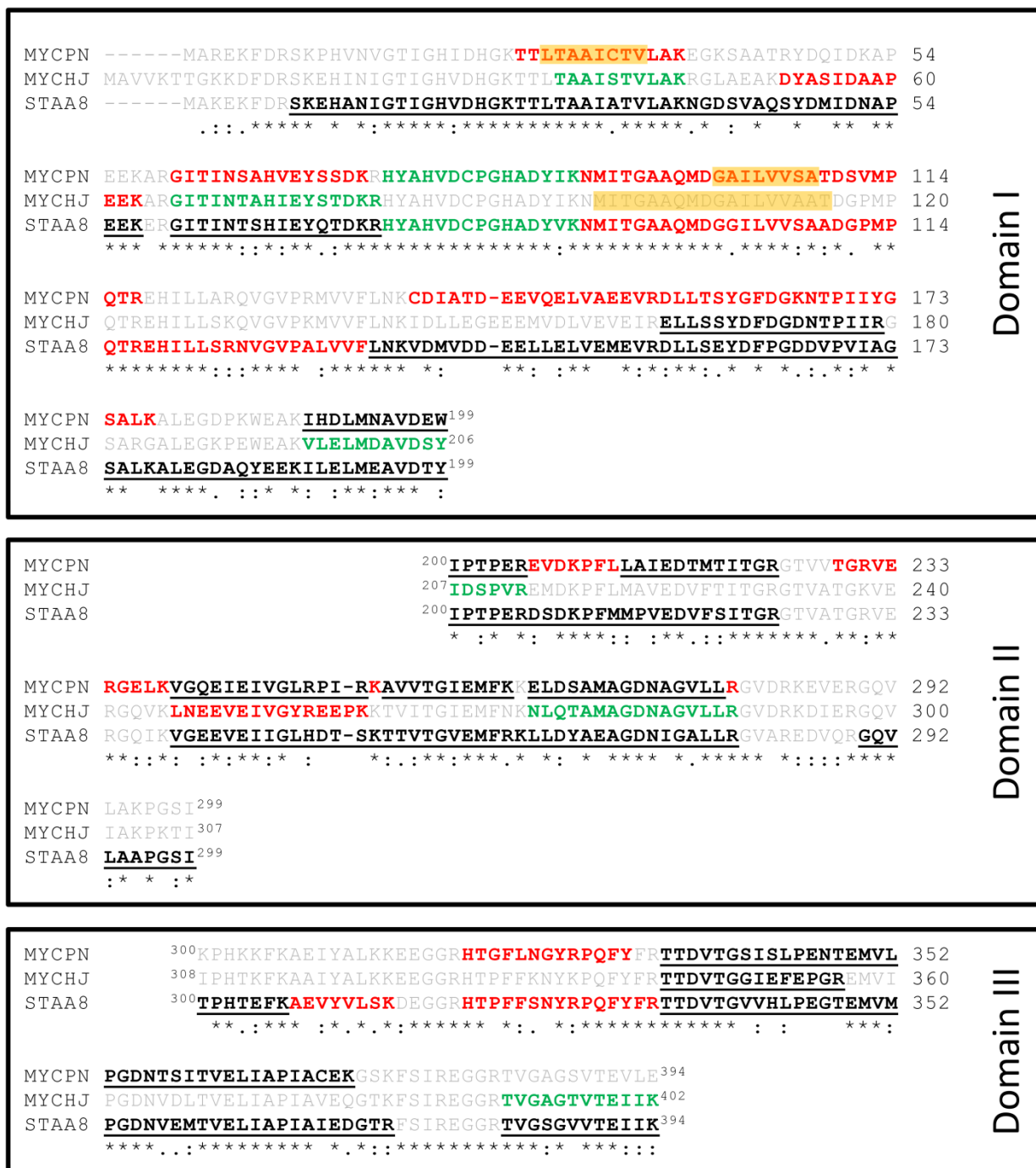
Prof. Steven P. Djordjevic

The itthree institute, University of Technology Sydney. PO Box 123, Broadway, NSW, 2007, Australia

Phone: +612 9514 4127

Fax: +612 9514 4143

Email: Steven.Djordjevic@uts.edu.au

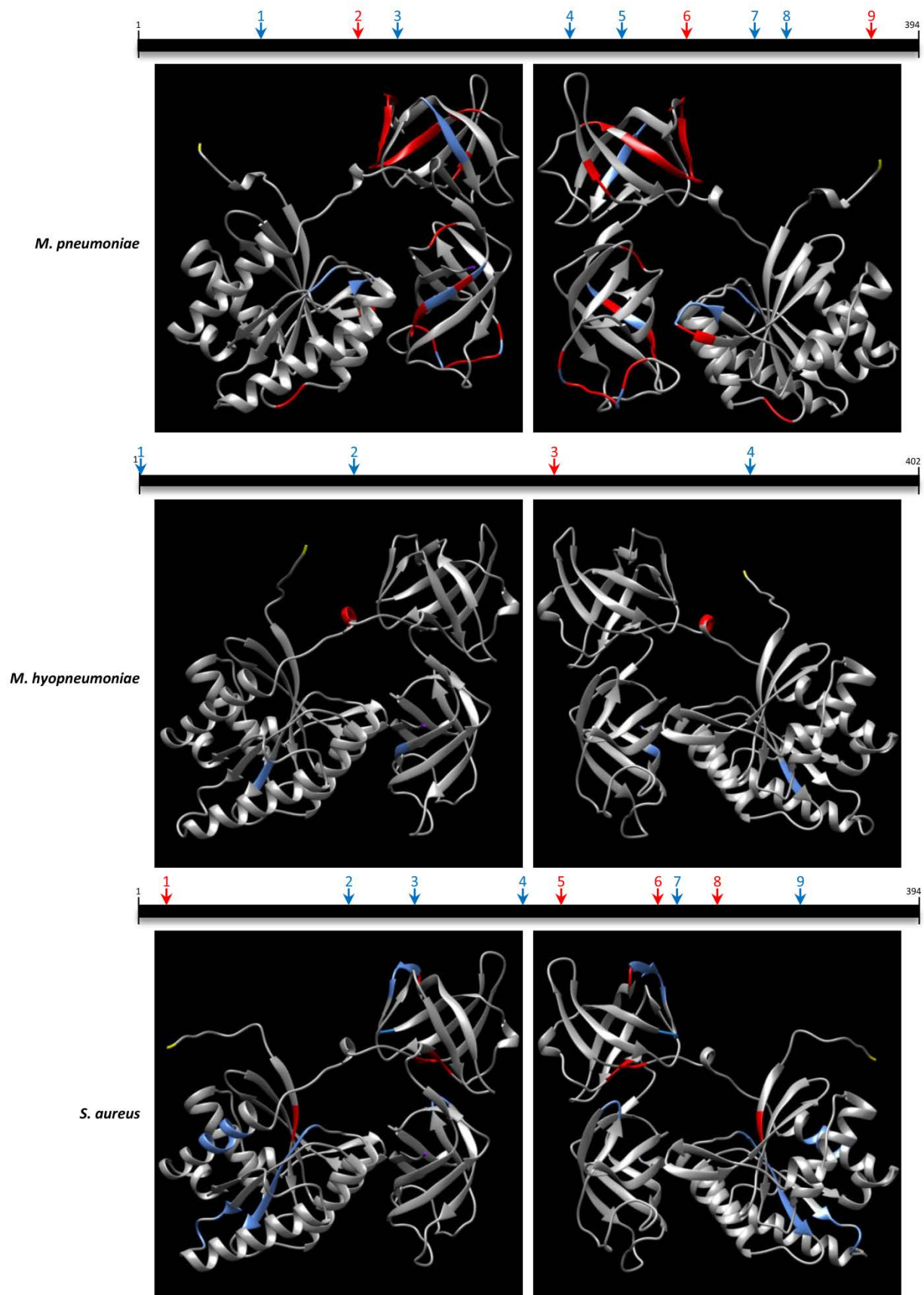
1065 **Supplementary Tables and Figures**

1066

1067 **Figure S1. Peptides that map to Ef-Tu identified from surface biotinylation and shaving**1068 **experiments. The sequence alignment for all three pathogens are shown separated into the three**

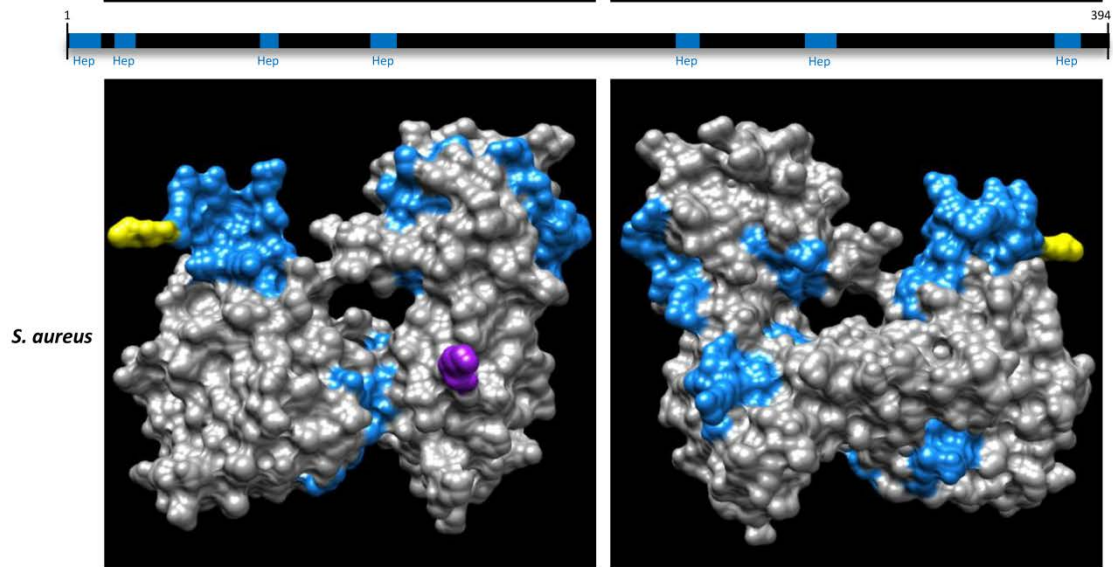
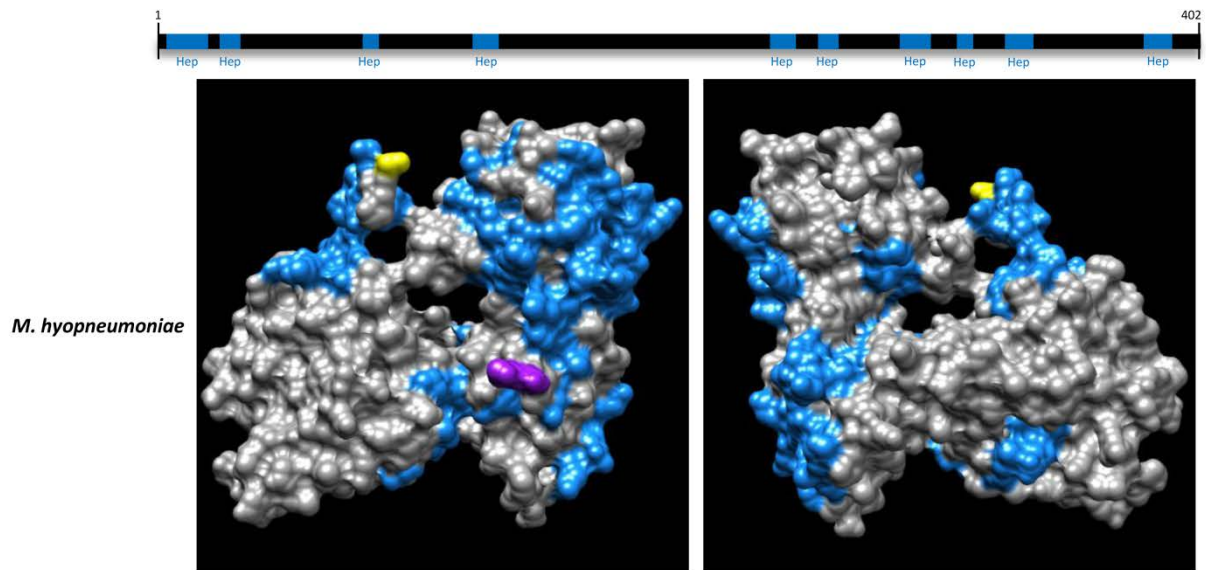
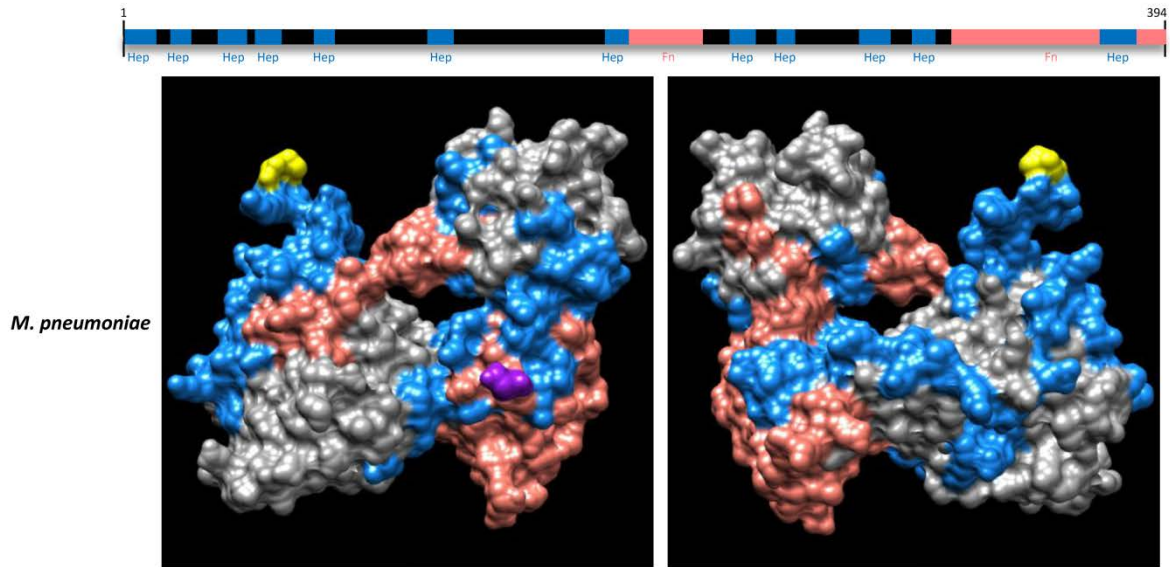
1069 domains of Ef-Tu. Peptides identified from surface trypsin shaving of whole cells are coloured
1070 green where red text indicate peptides from biotinylated surface proteins. Peptides that were
1071 identified from both experiments are highlighted as black underlined text. Yellow boxes indicate
1072 transmembrane domains for: *M. pneumoniae* adapted from ⁸⁹ and *M. hyopneumoniae* predicted
1073 by TMpred ¹⁴².

1074



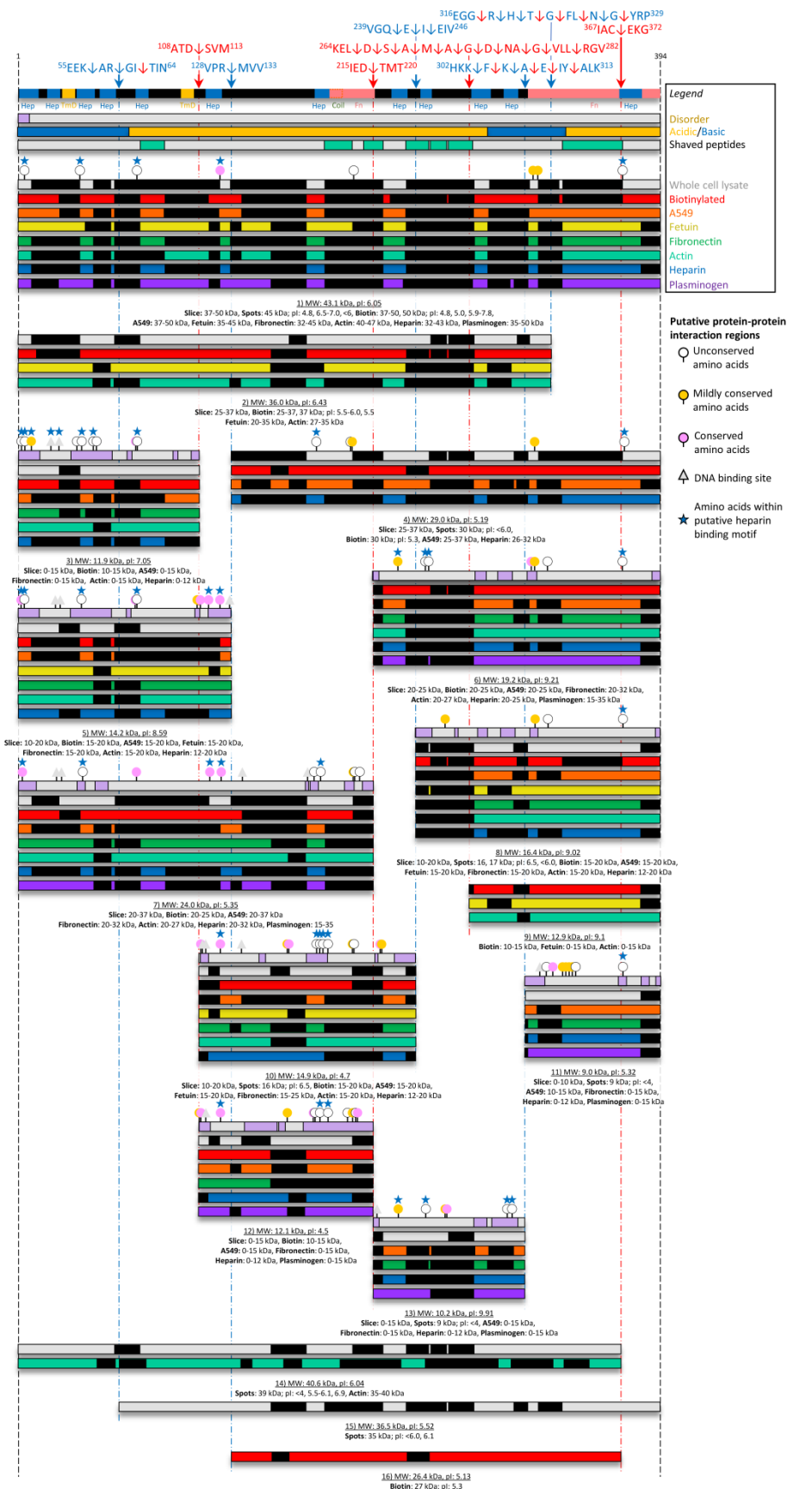
1076 **Figure S2. Predicted 3D ribbon structures of Mpn_{Ef-Tu}, Mhp_{Ef-Tu} and Sa_{Ef-Tu}.** Ef-Tu
1077 molecules are represented as black bars with identified cleavage sites displayed as arrows (blue,
1078 dimethyl labelling and red, semi-tryptic) with numbers to indicate each distinct cleavage site.
1079 Cleavage sites can also be seen in the ribbon structures as blue and red sections for dimethyl
1080 labelling and semi-tryptic sites, respectively. Images are reversed sides of each structure.
1081 Structures were predicted by MODELLER ⁹⁷ and were based on *E. coli* Ef-Tu homologues:
1082 PDB: 4G5G_A (*M. pneumoniae*), PDB: 1DG1_H (*M. hyopneumoniae*) and PDB: 1DG1_H (*S.*
1083 *aureus*). For context, the N-terminus of the protein is shown in yellow, and the C-terminus in
1084 purple.

1085



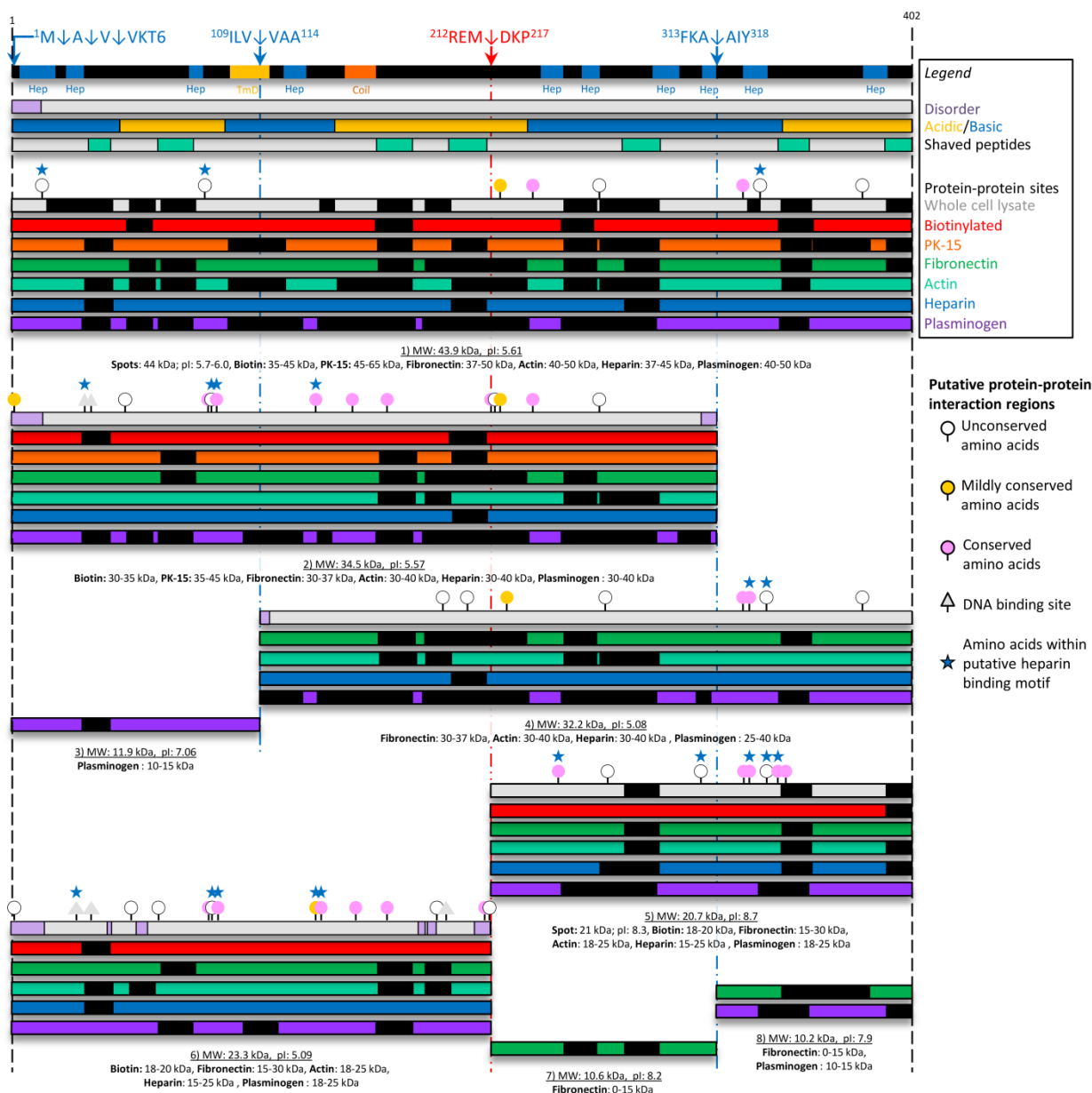
1087 **Figure S3. Predicted 3D space-filling structures for all three pathogens.** Full length Ef-Tu
1088 are represented as black bars with predicted heparin-binding sites for each pathogen shown in
1089 blue boxes and regions in the structures. The two published fibronectin-binding domains for *M.*
1090 *pneumoniae* have also been mapped. Structures were predicted by MODELLER and were based
1091 on *E. coli* Ef-Tu homologues: PDB: 4G5G_A (*M. pneumoniae*), PDB: 1DG1_H (*M.*
1092 *hyopneumoniae*) and PDB: 1DG1_H (*S. aureus*). The N-terminus of the protein is shown in
1093 yellow, and the C-terminus in purple.

1094



1096 **Figure S4. Cleavage map of Mpn_{Ef-Tu}.** Peptides (black boxes in coloured bars) identified by
1097 mass spectrometry of affinity chromatography of A549 surface proteins (orange bars), fetuin
1098 (yellow bars), fibronectin (green bars), actin (teal bars) and plasminogen (purple bars). Circles
1099 and triangles fragments indicate amino acid binding sites with either proteins or DNA,
1100 respectively. Purple boxes within grey bars are disordered regions within fragments. Red bars
1101 indicate peptides identified from surface biotinylation and peptides released from trypsin surface
1102 shaving can be seen as the green boxes in the grey bar above the full length protein.

1103



1104

1105 **Figure S5. Cleavage map of Mhp_{EF-Tu}.** Peptides (black boxes in coloured bars) identified by

1106 mass spectrometry of affinity chromatography of PK-15 surface proteins (orange bars),

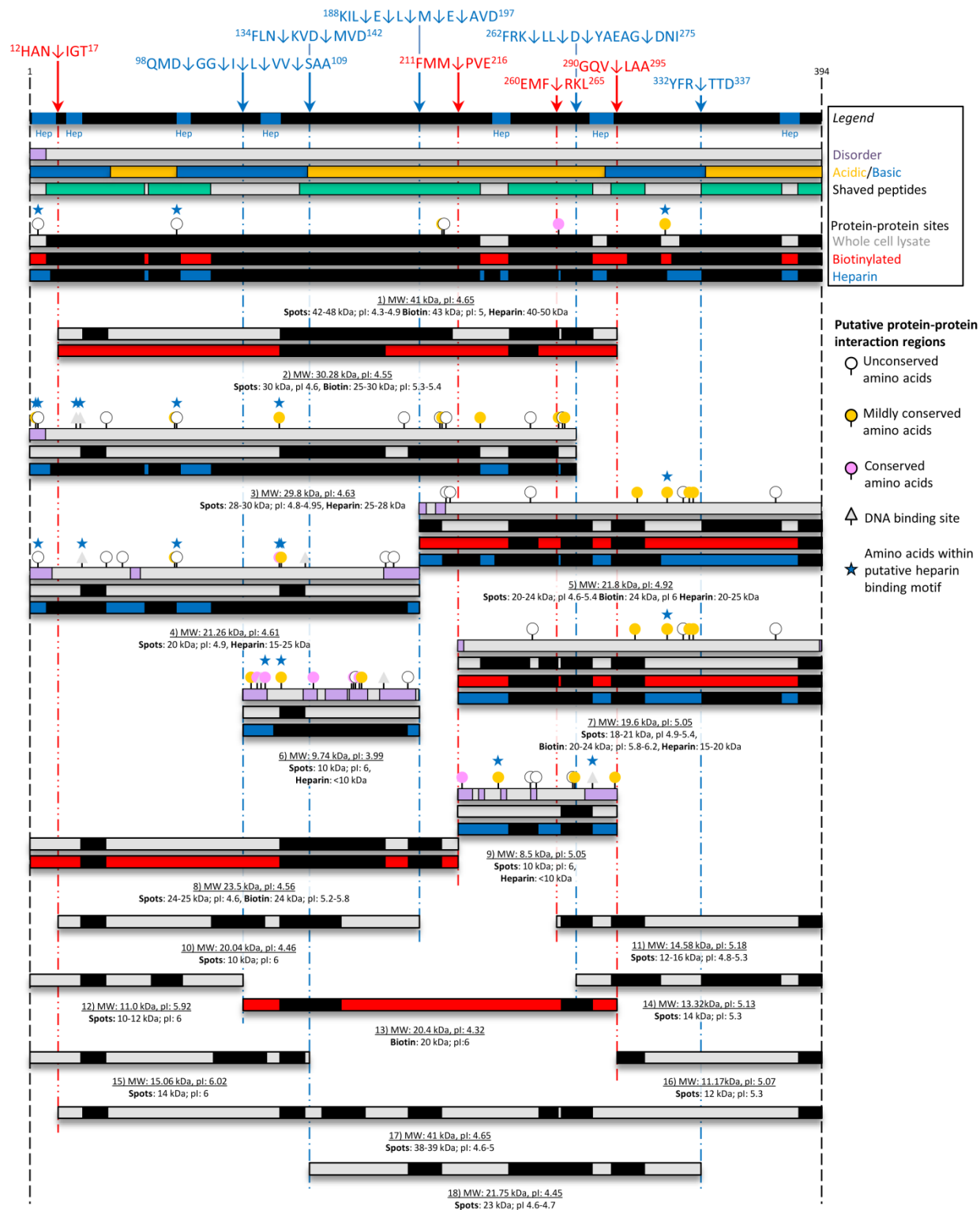
1107 fibronectin (green bars), actin (teal bars) and plasminogen (purple bars). Circles and triangles

1108 fragments indicate amino acid binding sites with either proteins or DNA, respectively. Purple

1109 boxes within grey bars are disordered regions within fragments. Red bars indicate peptides

1110 identified from surface biotinylation and peptides released from trypsin surface shaving can be
1111 seen as the green boxes in the grey bar above the full length protein.

1112



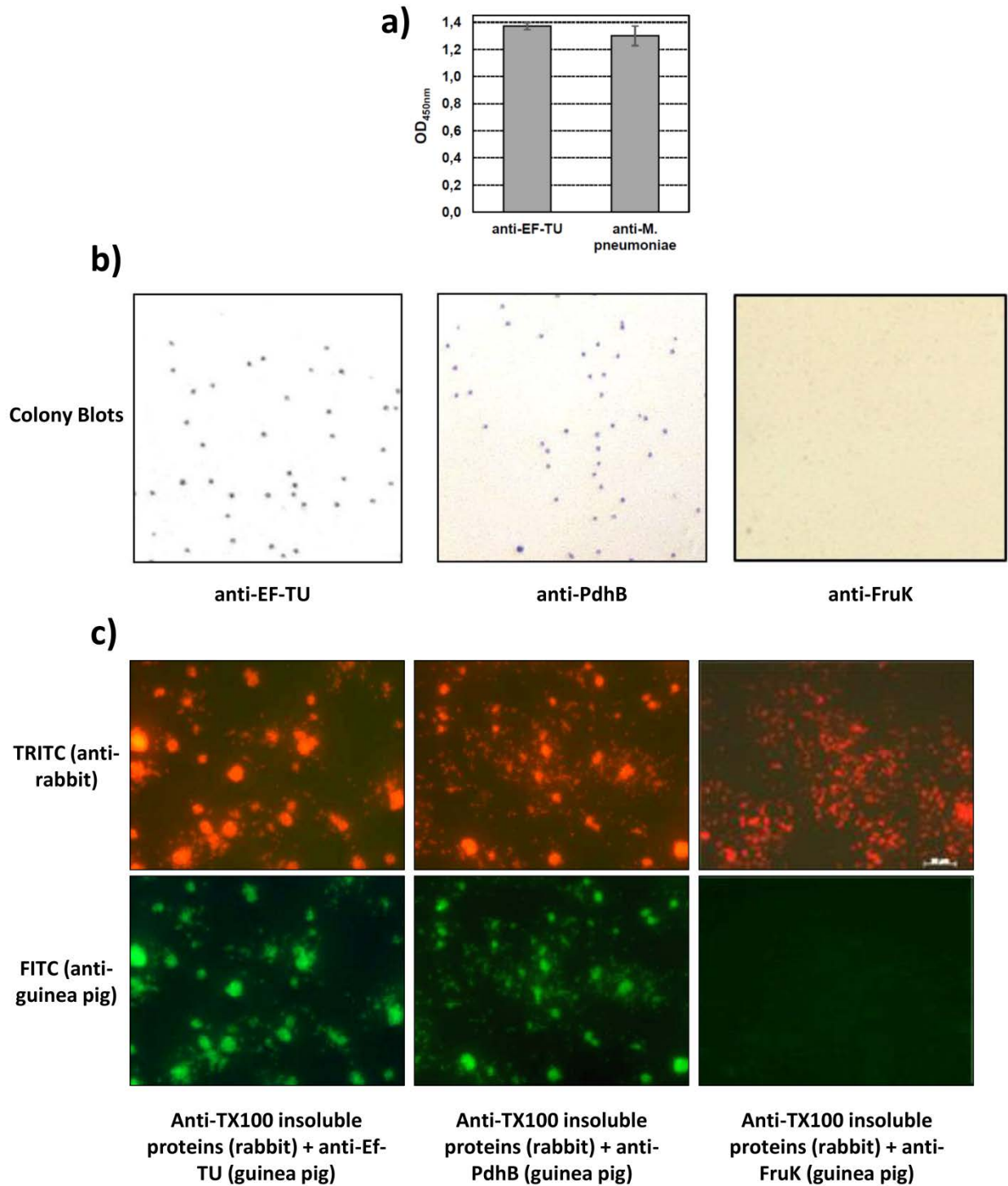
1113

1114 **Figure S6. Cleavage map of Ef-Tu_{Sa}.** An expansion of the cleavage maps in Figure 2. Includes

1115 peptides (black boxes in coloured bars) identified by mass spectrometry of Ef-Tu fragments that

1116 don't bind heparin or not on the surface (Fragments 10 – 18). Circles and triangles fragments
1117 indicate amino acid binding sites with either proteins or DNA, respectively. Purple boxes within
1118 grey bars are disordered regions within fragments. Red bars indicate peptides identified from
1119 surface biotinylation and peptides released from trypsin surface shaving can be seen as the green
1120 boxes in the grey bar above the full length protein.

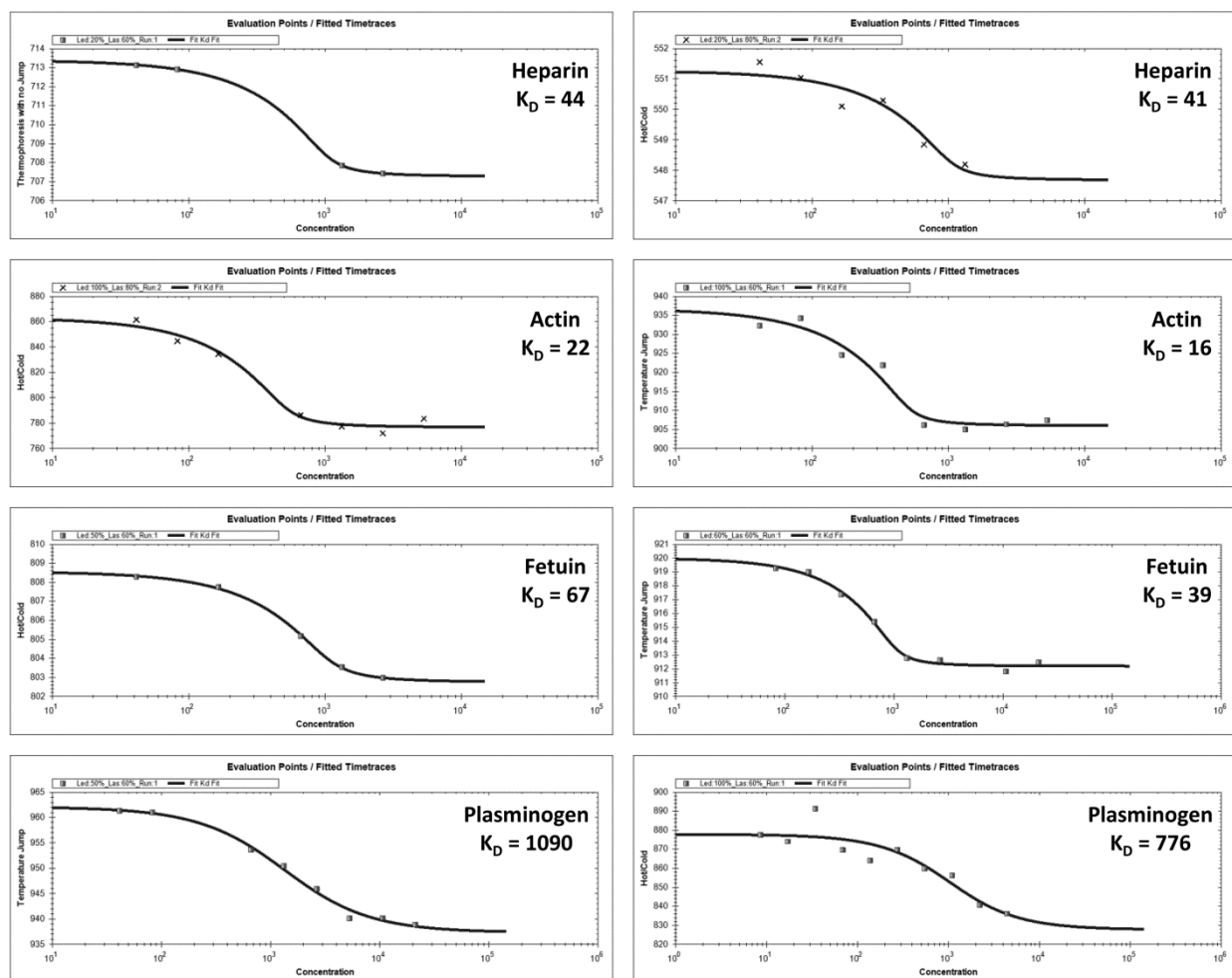
1121



1122

1123 **Figure S7: Mpn_{Ef-Tu} resides on the surface of *M. pneumoniae*.** A) Anti - rMpn_{Ef-Tu} antibodies1124 recognise *M. pneumoniae* cells in a whole cell ELISA assay; eight replicates. B) Colony blots of1125 *M. pneumoniae* probed with anti - rMpn_{Ef-Tu} antibodies, PdhB (positive control) and 1-

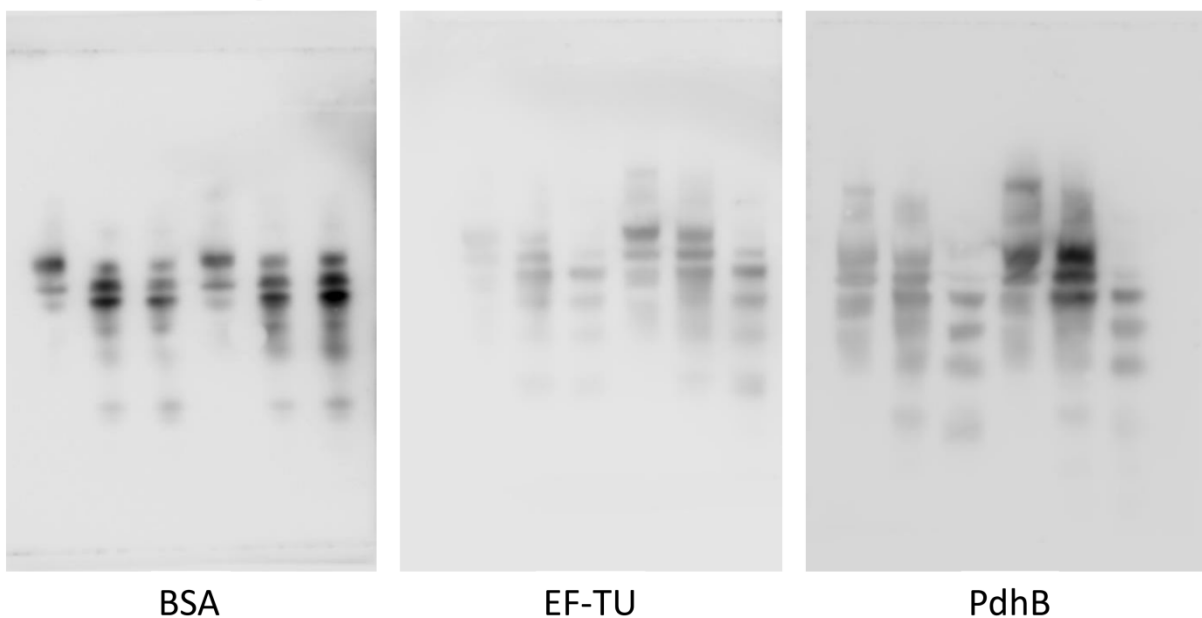
- 1126 phosphofructokinase (negative control). C) Immunofluorescence microscopy of *M. pneumoniae*
- 1127 cells probed with antibodies against TX100 insoluble proteins (TRITC; cell control), Ef-Tu,
- 1128 PdhB (positive control) and 1-phosphofructokinase as the negative control (FITC).
- 1129



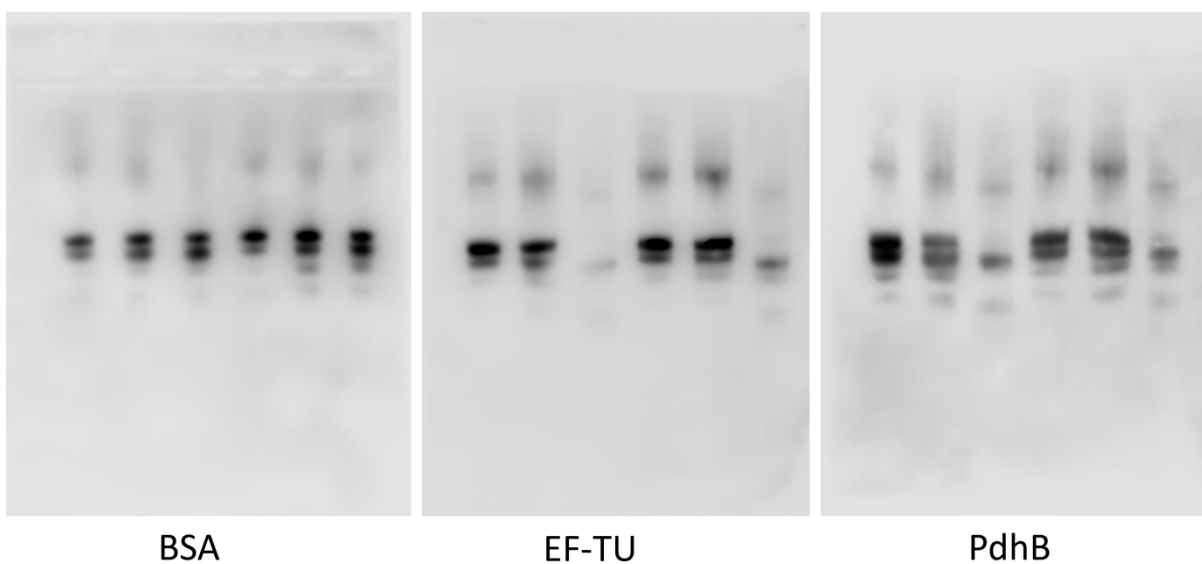
1130

1131 **Figure S8: Microscale thermophoresis output depicting the interaction of rMpn_{EF-Tu} with**
 1132 **human molecules.** Concentration of rMpn_{EF-Tu} is plotted against thermophoretic movement of
 1133 fluorescent human molecules. Experiments performed in duplicate, each panel representing one
 1134 replicate.

Human fibrinogen



Vitronectin



1135

1136 **Figure S9: Full length blots showing degradation of human fibrinogen and vitronectin.**

1137 Contrast adjusted (increased) to demonstrate no additional bands in full blots. Corresponding to

1138 Figure 5B, degradation with plasminogen activated by either urinary plasminogen activator or

1139 tissue plasminogen activator. Contrast and brightness editing are equal across the blots.

1140 Will be presented as “Supplementary file S8- Tables”

1141 **Table S1. Putative heparin-binding motifs identified in Mpn_{EF-Tu}, Mhp_{EF-Tu}, and Sa_{EF-Tu}.**

| Amino acid range | Sequence |
|---|-----------------------------|
| <u>Mpn_{EF-Tu} (Uniprot #: P23568)</u> | |
| 2 - 13: | aReKfdRsKpHv |
| 19 - 26: | gHidHgKt [≠] |
| 37 - 47: | aKegKsaatRy |
| 51 - 60: | dKapeeKaRg |
| 73 - 80: | dKRHyAHv* |
| 116 - 125: | tReHillaRq [≠] |
| 183 - 191: | pKweaKiHd |
| 230 - 239: | gRveRgelKv [≠] |
| 248 - 254: | lRpiRka |
| 279 - 290: | lRgvdrKeveRg |
| 299 - 307: | iKpHKKfKa |
| 370 - 383: | eKgsKfsiReggRt [≠] |
| <u>Mhp_{EF-Tu} (Uniprot #: Q4A9G1)</u> | |
| 4 - 19: | vKttgKKdfdRsKeHi |
| 25 - 32: | gHvdHgKt [≠] |
| 79 - 86: | dKRHyAHv* |
| 122 - 131: | tReHillsKq [≠] |
| 237 - 246: | gKveRgqvKl [≠] |
| 255 - 262: | yReepKKt |
| 287 - 298: | lRgvdrKdieRg |
| 302 - 315: | aKpKtiipHtKfKa |
| 327 - 337: | gRHtpffKnyKp |
| 381 - 391: | tKfsiReggRt [≠] |
| <u>Sa_{EF-Tu} (Uniprot #: Q2G0N0)</u> | |
| 2 - 13: | aKeKfdRsKeHa |
| 19 - 26: | gHvdHgKt [≠] |
| 73 - 80: | dKRHyAHv* |
| 116 - 125: | tReHillsRn [≠] |
| 230 - 239: | gRveRgqiKv [≠] |
| 279 - 290: | lRgvaRedvqRg |
| 373 - 383: | tRfsiReggRt [≠] |

1142 Mpn_{EF-Tu}, Mhp_{EF-Tu}, and Sa_{EF-Tu} were searched for patterns x-[HKR]-x(0,2)-[HKR]-x(0,2)-[HKR]-x and x-

1143 [HKR]-x(1,4)-[HKR]-x(1,4)-[HKR]-x using ScanProsite¹⁴⁴. * indicates the motif (dKRHyAH) which is

1144 found in all three pathogens. [≠] indicates the motifs that are highly homologous (up to three non-basic

1145 residues different) in all three pathogens.

1146

1147 **Table S2A. Number of binding sites in full length and fragments of Mpn_{EF-Tu}.**

1148 Analysis of the Mpn_{EF-Tu} for the putative protein-protein (P:P) and protein-nucleic acid
 1149 interaction sites using ISIS ⁹⁵.

1150 **Mpn_{EF-Tu} (Uniprot #: P23568)**

| Fragment number | Full length | 3 | 4 | 5 | 6 | 7 | 8 | 10 | 11 | 12 | 13 |
|--------------------|-------------|-------|---------|-------|---------|-------|---------|---------|---------|---------|---------|
| Range (amino acid) | 1-394 | 1-110 | 131-394 | 1-130 | 218-394 | 1-217 | 244-394 | 111-242 | 308-394 | 111-217 | 218-307 |
| Exposed P:P | 8 | 15 | 5 | 13 | 10 | 17 | 6 | 12 | 9 | 19 | 11 |
| Buried P:P | 6 | 5 | 2 | 7 | 5 | 2 | 5 | 3 | 6 | 2 | 0 |
| DNA-binding | 0 | 8 | 0 | 5 | 0 | 6 | 0 | 3 | 2 | 1 | 1 |

1151

1152 **Table S2B. Number of binding sites in full length and fragments of Mhp_{EF-Tu}.**

1153 Analysis of the Mhp_{EF-Tu} for the putative protein-protein (P:P) and protein-nucleic acid
 1154 interaction sites using ISIS ⁹⁵.

1155 **Mhp_{EF-Tu} (Uniprot #: Q4A9G1)**

| Fragment number | Full length | 1 | 4 | 5 | 6 |
|--------------------|-------------|-------|---------|---------|-------|
| Range (amino acid) | Full | 1-315 | 112-402 | 215-402 | 1-215 |
| Exposed P:P | 10 | 13 | 9 | 10 | 18 |
| Buried P:P | 2 | 3 | 1 | 3 | 4 |
| DNA-binding | 0 | 3 | 0 | 0 | 4 |

1156

1157 **Table S2C. Number of binding sites in full length and fragments of Sa_{EF-Tu}.**

1158 Analysis of the Sa_{EF-Tu} for the putative protein-protein (P:P) and protein-nucleic acid interaction
 1159 sites using ISIS ⁹⁵.

1160 **Sa_{EF-Tu} (Uniprot #: Q2G0N0)**

| Fragment number | Full length | 3 | 4 | 5 | 6 | 7 | 9 |
|---------------------------|--------------------|--------------|--------------|----------------|----------------|----------------|----------------|
| Range (amino acid) | Full | 1-266 | 1-192 | 193-394 | 104-192 | 214-394 | 214-292 |
| Exposed P:P | 10 | 13 | 10 | 7 | 12 | 5 | 10 |
| Buried P:P | 6 | 12 | 6 | 5 | 2 | 5 | 1 |
| DNA-binding | 0 | 3 | 4 | 0 | 2 | 0 | 1 |

1161

1162

1163 Will be presented as “Supplementary file S9- Bioinformatics”

1164 **Supplementary Materials - Bioinformatics**

1165 **Conserved/Non-conserved regions:**

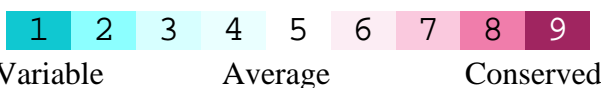
1166 **Analysis of conservation of amino acids in Mpn_{EF-Tu}, Mhp_{EF-Tu}, and Sa_{EF-Tu}.**

1167 Degree of conservation in Mpn_{EF-Tu} was calculated by The ConSurf server⁹⁶. Colours indicate
 1168 degree of conservation of the amino acid across species. Red diamond sticks indicate predicted
 1169 protein binding sites by ISIS^{95,149}. Grey, yellow and purple circle sticks indicate predicted
 1170 nucleotide, DNA and RNA binding regions, respectively by SomeNA^{149,150}. Below the sticks are
 1171 three rows: i) Blue and red bars indicate predicted beta-strand and helix secondary structures,
 1172 respectively by REPROFSec^{147,149}, ii) Blue and yellow bars indicate predicted regions that are
 1173 exposed and buried to solvent accessibilities, respectively by PROFAcc^{147,149}. iii) Green bars
 1174 indicate predicted disordered regions by Meta-Disorder (MD)^{146,149}. Circled amino acids
 1175 represent predicted binding sites. Amino acids indicated by a triangle indicate predicted
 1176 nucleotide binding sites.

1177 **Legend:**

1178

1179 The conservation scale:



1180 **e** - An exposed residue according to the neural-network algorithm.

1181 **b** - A buried residue according to the neural-network algorithm.

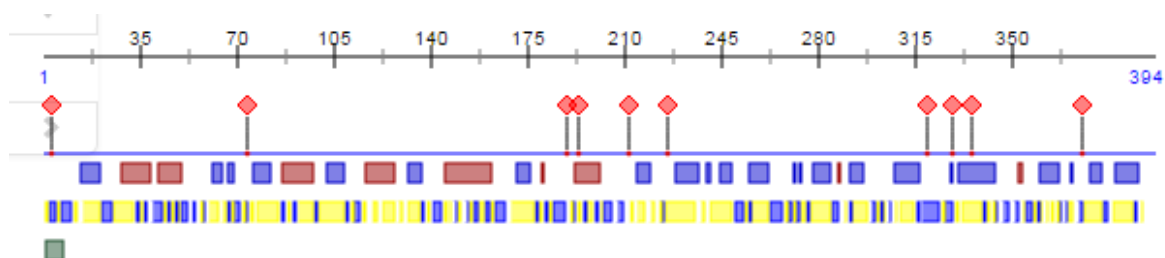
1182 **f** - A predicted functional residue (highly conserved and exposed).

1183 **s** - A predicted structural residue (highly conserved and buried).

1184 **x** - Insufficient data - the calculation for this site was performed on less than 10% of
 1185 the sequences.

1186 **Mpn_{Ef-Tu} (Uniprot #: P23568)**

1187 **Full length protein: 1-394**



1188

| | | | | |
|--|--------------------------|--------------------------|--|------------|
| 1 | 11 | 21 | 31 | 41 |
| MAREKFD ¹ RSK | PHVNVGTIGH | IDHGKTTLTA | AICTVLAKEG | KSAATRYDQI |
| 51 | 61 | 71 | 81 | 91 |
| DKAPEEKARG | ITINSAHVEY | SSDKRHYAHV | DCPGHADYIK | NMITGAAQMD |
| 101 | 111 | 121 | 131 | 141 |
| GAILVVSATD | SVMPQ ² TREHI | LLAROVGVPR | MVVF ³ L ⁴ NKCDI | ATDEEVQELV |
| 151 | 161 | 171 | 181 | 191 |
| AE ⁵ EV ⁶ RDLLTS | YGF ⁷ DGKNTPI | IYGSALKALE | GDPKWEAKIH | DLMNAVDEWI |
| 201 | 211 | 221 | 231 | 241 |
| PTPEREVDKP | FLLAIEDTMT | ITGRGTVVTG | RVERGELKVG | QEIEIVGLRP |
| 251 | 261 | 271 | 281 | 291 |
| IRKAVVTGIE | MEKKELDSAM | AGDNAGVILR | GVDRKEVERG | QVLAKPGSIK |
| 301 | 311 | 321 | 331 | 341 |
| PHKKEKAEIY | AIKKE ⁸ EGGRH | TGFLNGYRPQ | FYFRTTDVTG | SISLPENTEM |
| 351 | 361 | 371 | 381 | 391 |
| VLPGDNTSIT | VELIAPIACE | KGSKE ⁹ SIREG | GRTVGAGSVT | EVLE |

1189

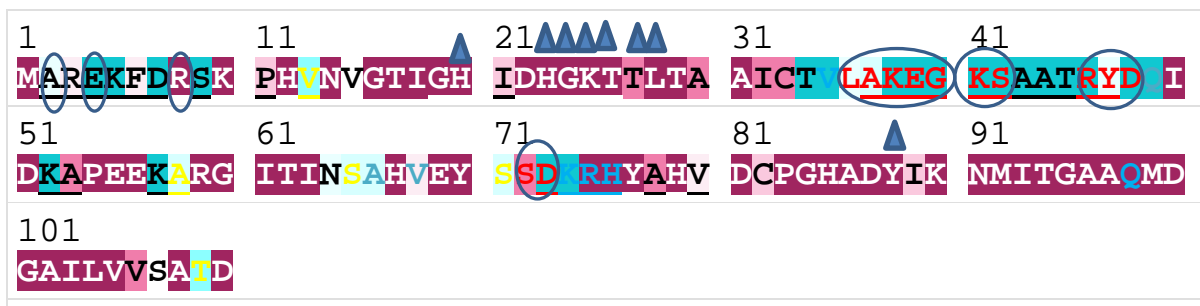
1190

1191 **Fragment 3: 1-110**

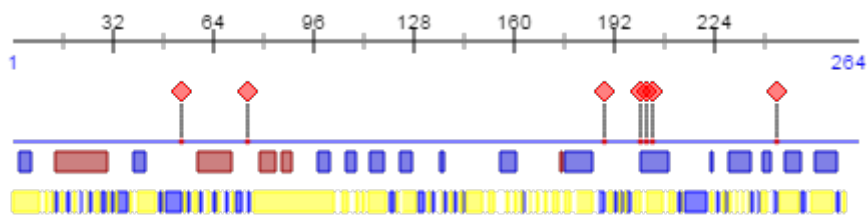


1192

1193



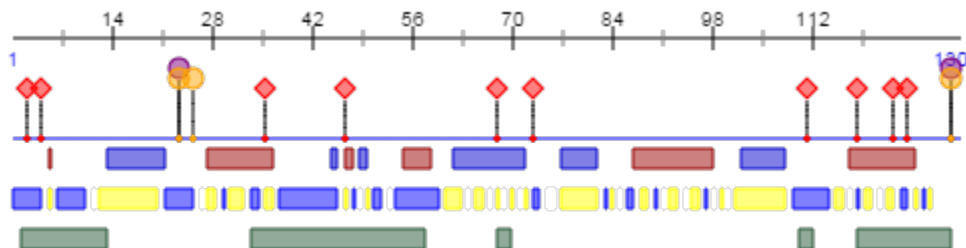
1194

1195 **Fragment 4: 131-394**

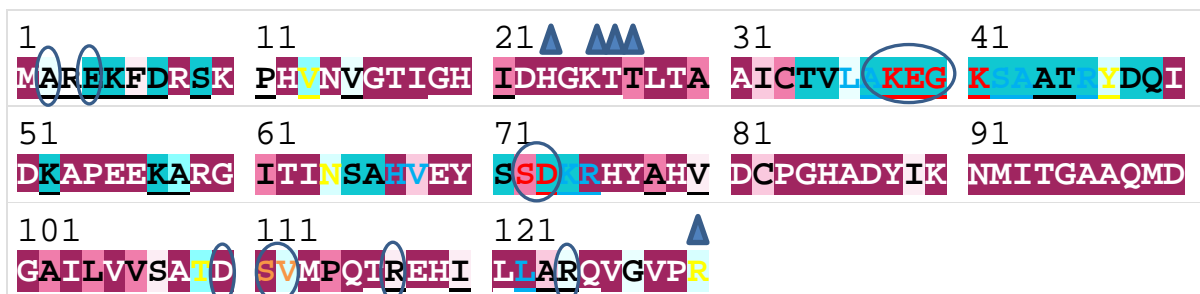
1196



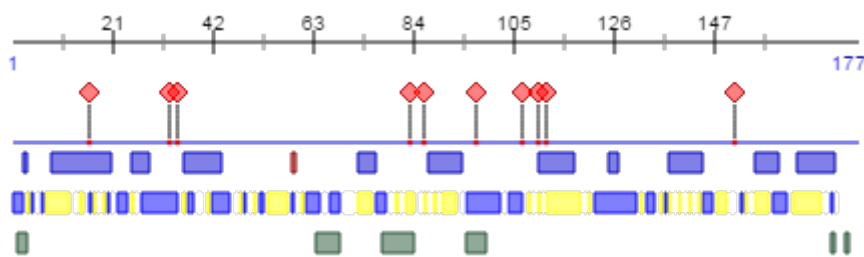
1197

1198 **Fragment 5: 1-130**

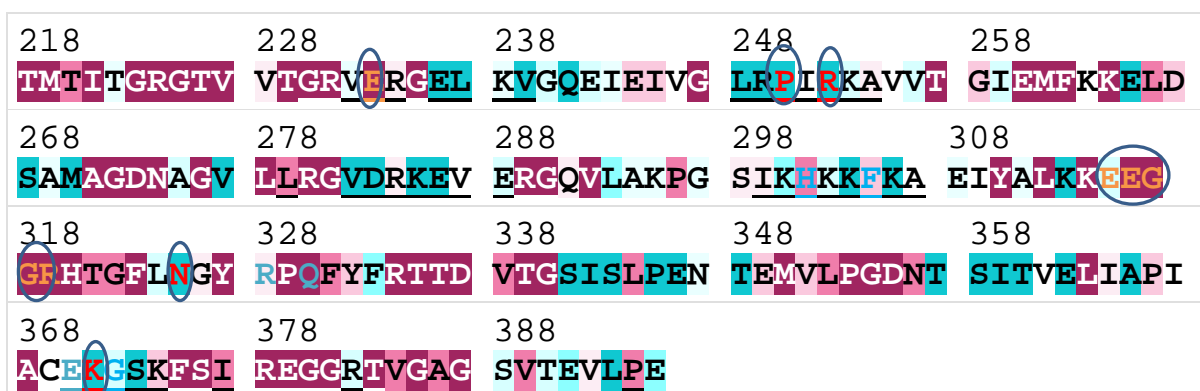
1199



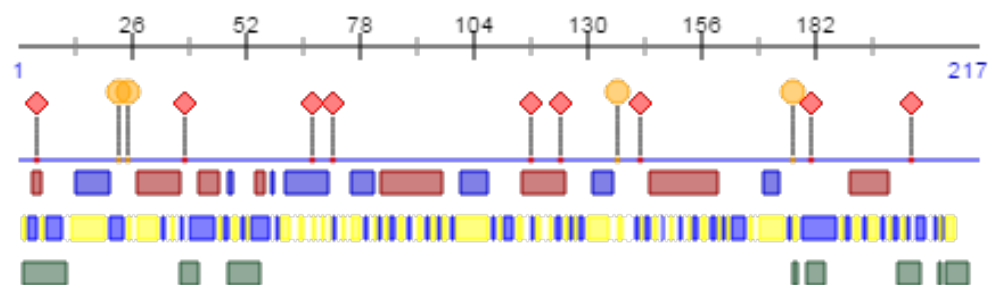
1200

1201 **Fragment 6: 218-394**

1202



1203

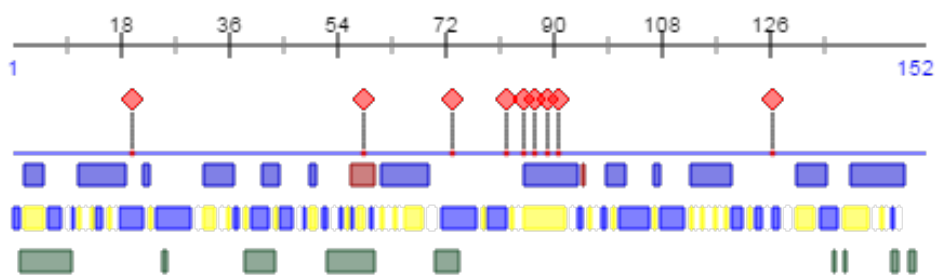
1204 **Fragment 7: 1-217**

1205



| | | | | |
|------------|------------|-------------|------------|-------------|
| 51 | 61 | 71 | 81 | 91 |
| DKAPEEKARG | ITINSAHVEY | SRRHYAHV | DCPGHADYIK | NMITGAAQMD |
| 101 | 111 | 121 | 131 | 141 |
| GAILVVSATD | SVMPQIREHI | LLARQVGVPR | MVVFLNKCDI | ATDEEVQELV |
| 151 | 161 | 171 | 181 | 191 |
| AAEVRDLLTS | YGFDGKNTPI | IYGSAALKALE | GDPKWEAKIH | DLMNAVDDEWI |
| 201 | 211 | | | |
| PTPREDKP | FLLAIED | | | |

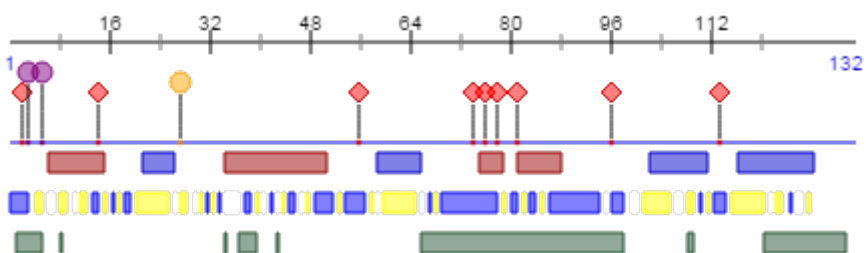
1206

1207 **Fragment 8: 244-394**

1208

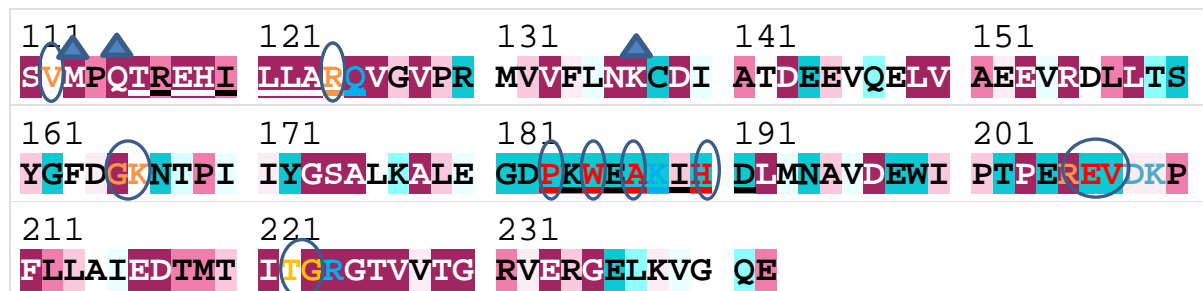
| | | | | |
|------------|-------------|------------|------------|------------|
| 243 | 253 | 263 | 273 | 283 |
| IEIVGLRPIR | KAVVTGIEMF | KKELDSAMAG | DNAGVLLRGV | DRKEVERGQV |
| 293 | 303 | 313 | 323 | 333 |
| LAKPGSIKPH | KKFKAETIYAL | KKEEGGRHTG | FLNGYRPOFY | FRTTDVTGSI |
| 343 | 353 | 363 | 373 | 383 |
| SLPENTEMVL | PGDNTSITVE | LIAPIACEKG | SKFSIREGGR | TVCAGSVTEV |
| 393 | | | | |
| LE | | | | |

1209

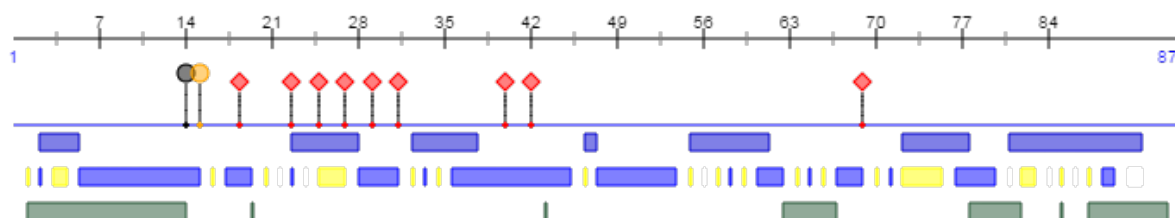
1210 **Fragment 10: 111-242**

1211

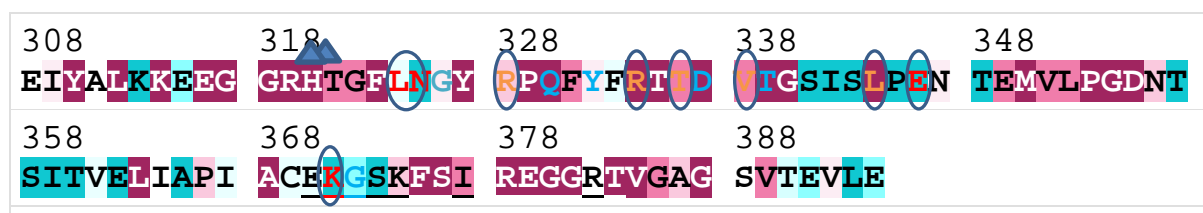
1212



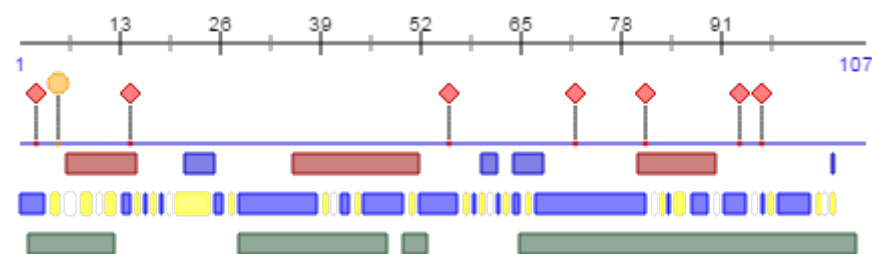
1213

1214 **Fragment 11: 308-391**

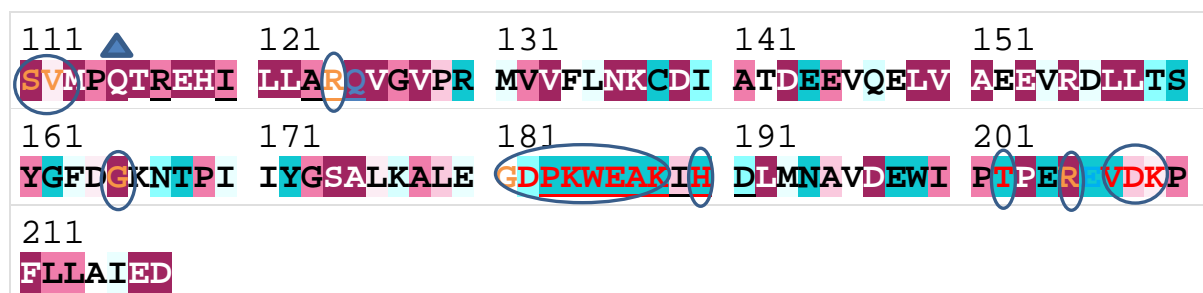
1215



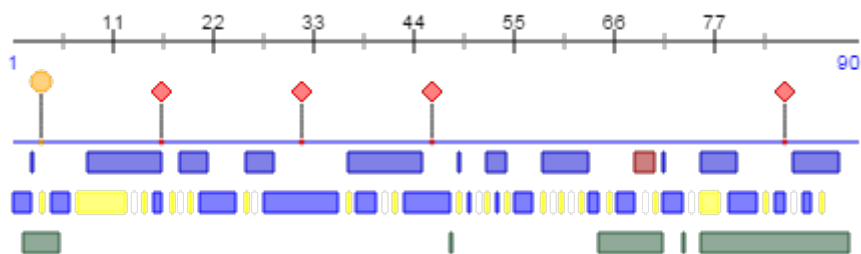
1216

1217 **Fragment 12: 111-217**

1218



1219

1220 **Fragment 13: 218-307**

1221

| | | | | |
|-------------------|--------------------------|---------------------------|--------------------------|--------------------------|
| 218 | 228 | 238 | 248 | 258 |
| TMTITGRGTV | VTGRVERGEL | KVGQEEIEIVG | LRPIRKAVVT | GIEMFKEELD |
| 268 | 278 | 288 | 298 | |
| SAMAGDNAGV | LLRGVDRKEV | ERGQVLAKPG | SIKPHKKFKA | |

1222

1223

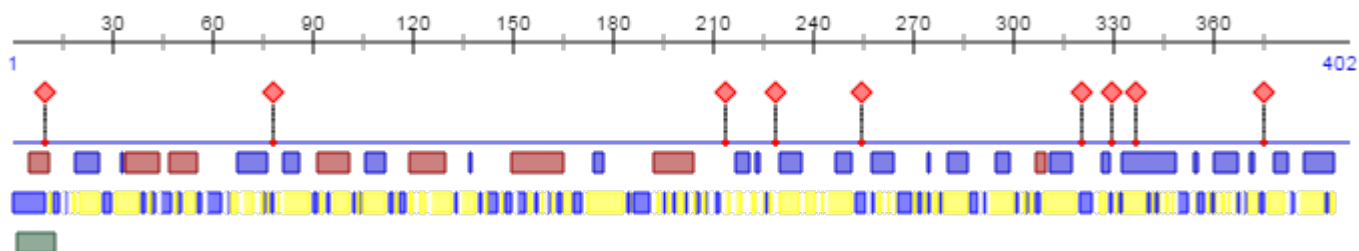
1224

1225

1226 Mhp_{Ef-Tu} (Uniprot #: Q4A9G1)1227 **Full length protein: 1-402**

1228

1229

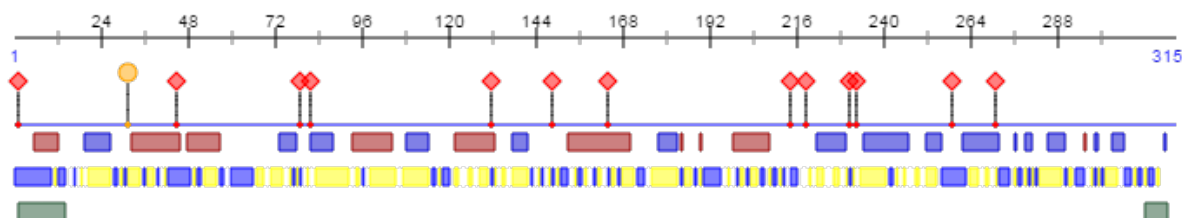


1230

| | | | | |
|-------------|-------------|-------------|-------------|-------------|
| 1 | 11 | 21 | 31 | 41 |
| MAVVKTIGKK | DFDRSKEHIN | IGTIGHVDHG | KTTLTAAIST | VLAKRGLAEA |
| eebeeebbbe | ebeeeeeebb | bbbbbbbeeb | ebbebobbbe | bbbeeeeeeb |
| s f | f f f s | sssss ffs | fsfsf ss | |
| 51 | 61 | 71 | 81 | 91 |
| KDYASIDAAP | EKARGITIN | TAHIEYSTDK | RHYAHVDCPG | HADYIKNMIT |
| eebeebbebe | eeeeebbbb | bbebebeeee | ebbbbbbbbe | ebebbbebbb |
| sf sf | ff fssss | f fs f | fssss s ff | fsfs ffsss |
| 101 | 111 | 121 | 131 | 141 |
| GAAQMDGAIL | VVAATDGPMP | QIREHILLSK | QVGVPKMVVF | LNKIDLLEGE |
| bbbbbebbbb | bbbbeeeee | ebbebbbbbe | ebbebbbbbb | bbbeeeeeee |
| sssssfsss | s s fffff | fsffs ss | fs sf s | sf f f f |
| 151 | 161 | 171 | 181 | 191 |
| EEMVDLVEVE | IRELLSSYDF | DGDNTPPIRG | SARGALEGKP | EWEAKVLEIM |
| eebbbebeee | beebbeeeeb | eeeebebbbe | bbbebebeee | eeeebeebbb |
| f s s f | f ss f | fff f s | ss | s |
| 201 | 211 | 221 | 231 | 241 |
| DAVDSYIDSP | VREMDKPFLLM | AVEDVFTITG | RGTVATGKVE | RGQVKLNEEV |
| eebeeebeee | eeeeeeebbb | ebbebbbbbb | ebbbebebe | ebbebebeeb |
| f f f | f fs s | ff s s | fsfs fsf f | fs |
| 251 | 261 | 271 | 281 | 291 |
| EIVGYREEPK | KTVITGIEMF | NKNLQTAMAG | DNAGVLLRGV | DRKDIERGOV |
| ebbbbeeeee | eebebbbebb | eebbebeeee | eebebbbebb | eeeebeebbe |
| f s | f fss | f sf ff | ff f s s fs | f fs s |
| 301 | 311 | 321 | 331 | 341 |
| IAPKTIIPH | TKFKAAIYAL | KKEEGRHTP | FFKNYKPQFY | FRTTVDVTGGI |
| beeeeeeeeee | eebebebbbbb | eeeeeeeeeee | ebeeeeeebb | bebeeeebbe |
| f ff | s s | f fffffff | f ffffs | fsffffs |

351 EFEPG**RE**MVI PGDNVDL**TV**E LI**APIA**VE**Q**S TKFSI**REGGR** TV**GAGT**VTEI
 ebeebbbbbb eeeeebebe bbbbbbbbeb bebbbbeeee bbbbbbebeb
 s s s ffff s s s sssffff sss s
 401
 IK
 ee

1231

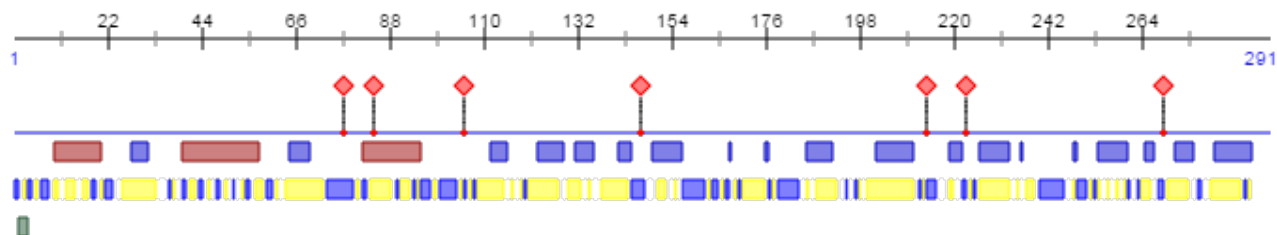
1232 **Fragment 1: 1-315**

1233

1 MA**VVK**TG**KK** DFDR**SKEH**IN IGT**IGHV**DHG K**TTLT**AA**IST** VLA**KR**GL**AEA**
 eebeeebbbe ebeeeeeebb bbbbbbbbeb ebebebbbbe bbbbeeeeeeb
 sssf f f f s sssss ffs fsfsfsss
 51 KD**YASID**AAP EEK**ARGIT**IN TA**HIEYS**TDK R**HYAHV**DCPG HAD**YIKN**MIT
 eebeebbebe eeeeebbbbb bbebebeeee ebbbbbbbee ebebbbebbb
 sf sf ff fssss f fs f fssss s ff fsfs ffsss
 101 GAA**QMDG**A**IL** VVA**ATDGP**M**P** Q**TREHIL**LSK R**VGVPK**M**VVF** L**NKIDL**EGE
 bbbbbbebbb bbbbbbeeee ebeebbbbbe ebebebbbbe bbebeeeeeee
 sssssfs s s s ff ff fsffs ss f fs sf s sf f f f
 151 E**EMVDL**VEVE I**RELLS**SYDF DGD**NTPI**IRG S**ARGAL**EG**KP** E**WEAKV**LE**IM**
 eebbebeeee beebbeeeeb eeeebbbbeb bbbbebeeee ebeeebeebb
 f sfff f ss f ff f ss
 201 DAV**SYID**SP V**REMDK**P**FLM** AV**EDVFT**ITG R**GTVA**TG**KVE** R**GOVKI**NE**EF**
 ebbeeebeee eeeeeebbbb ebeebbbbbe ebebbbebe ebebeeeeeeb
 f f f f fs s ff s s fsfs fsf f fs
 251 EIV**GYR**EEP**K** K**TVITG**IE**MF** N**KNLQ**TAMAG D**NAGV**LL**RGV** D**RKDI**ER**GQV**
 ebbbbbeeee eebebbbeb eebbebeeee eebbbbeb eeebebebeb

s f fss f sf ff ff f ssfs fs s
 301 311
 IAKPKTIIPH TKFKA
 beeeeeeeee eeeee
 f f f f

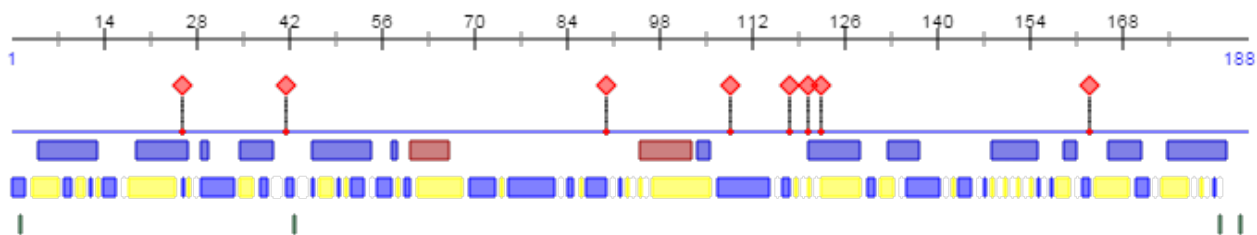
1234

1235 **Fragment 4: 112-402**

1236

112 122 132 142 152
 VAATDGPMPO TREHILLSKO VGVPKMVFL NKIDLLEGE EMVDLVEVEI
 eeeeeeeeee beebbbbbee beebbbbbb beeeeeeeee ebbbeeeeb
 f f ffffff s fs ss f s sf s sf f f f f ssf f
 162 172 182 192 202
 RELSSYDFD GDNTPIIRGS ARGALEGKPE WEAQVLEIMD AVDSYIDSPV
 eebbeeebe eeebebbb bbbebeeee beeebebbe bbeeebeeee
 f ss ff f s s s s f f f
 212 222 232 242 252
 REMDKPFLMA VEDVFTITGR GTVATGKVER GOVKLNEVE IVGYREPKK
 eeeeebbbe beebbbbbe bebbebebe beebbeeee bbbbeeeeee
 f fs s ffs s sf sfs fsf ff s f s
 262 272 282 292 302
 TVITGIEMFN KNLQTAMAGD NAGVLLRGVD RKDIERGQVI AKPKTIIPH
 eebbebbe ebeeeeee ebbbbebbe eebbeebbb eeeeeeeee
 f fss f sf fff f s ssfs f fs s f ff
 312 322 332 342 352
 KFKAAYALK KE~~EGGR~~TPF FKNYKPOFYF RTTDVTGGIE FEPGREMVIP
 ebeebbbbbe eeeeeeeee beeeeeebb ebeeebebe beebbebbb
 s s s f ffffffff ffffs fsffffs s fs sf
 362 372 382 392 402
 GDNVDLTVEL IAPIAVE~~Q~~GT KESIREGCRV VGACTVTEII K
 eebbebebeb bbbbbbbeeb ebbbeeeeb bbbbebbebe e
 fff s s s s sffff s ss s

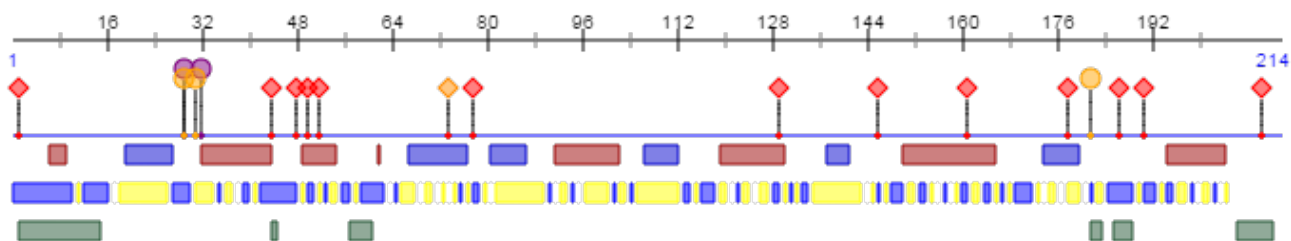
1237

1238 **Fragment 5: 215-402**

1239

| | | | | |
|------------|------------|------------|------------|-------------|
| 215 | 225 | 235 | 245 | 255 |
| DKPFLMAVED | VFTITGRGTV | ATGKVERGOV | KLNEEVEIVG | YREEPKKTVI |
| eeeeebabee | bbbbbbebeb | bebebeeb | eeeeebbbb | beeeeeeeeb |
| f ff sf ff | s s sfsfs | fsf ffs | f s | |
| 265 | 275 | 285 | 295 | 305 |
| TGIEMFNKNL | QTAMAGDNAG | VLLRGVDRKD | IERGOVIAKP | KTIIPHTKFK |
| ebbbebeebb | eebeeeebbb | bbbebeeb | bebebeeb | eeeeeeeb |
| f fssff | f ffff s | ssfs | fs s f | f |
| 315 | 325 | 335 | 345 | 355 |
| AAIYALKKEE | GGHTPFEKN | YKPOFYFRTT | DVTGGIEFEP | GREMVI PGDN |
| bebbbbeeee | eeeeeeeb | eeebbbebe | eebebebe | bbbbbbeeee |
| s s f f | ffffff | ffffs fsf | fffs s | s fs sffff |
| 365 | 375 | 385 | 395 | |
| VDLTVELIAP | IAVEQGTKFS | IREGGRIVGA | GTVTEIIK | |
| bebebebbb | bbbeebbeb | beeeebbbb | bbbbebee | |
| s | s s s | fffffsss | s s | |

1240

1241 **Fragment 6: 1-215**

1242

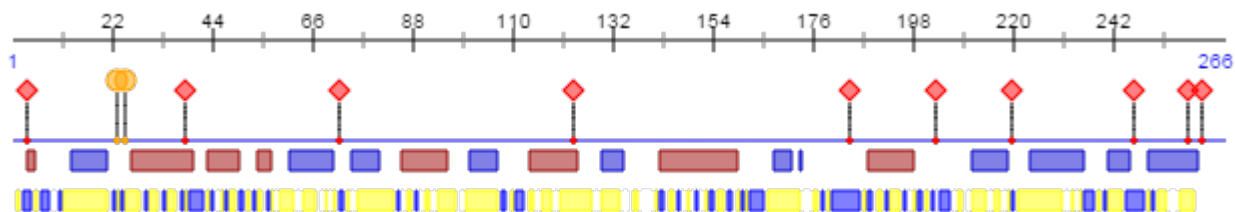
| | | | | |
|------------|-------------|------------|------------|------------|
| 1 | 11 | 21 | 31 | 41 |
| MAVVKTIQKK | DFDRSKEHIN | IGTIGHVDHG | KTTLTAAIST | VLA KRGLAE |
| eebeeebb | ebeeeeeb | bbbbbbeeb | ebbebbbbb | bbbeeeeb |
| s f | f f f s | sssss ffs | fsfsfsss | |
| 51 | 61 | 71 | 81 | 91 |
| KDYASIDAAP | E EKARGITIN | TAHIEYSTDK | HYAHVDCPG | HADYIKNMIT |
| eebeebbee | eeeeebbbb | bbbebeeee | ebbbbbbb | ebbbeebbb |

sf sf ff fssss f fs f fssss s ff fsfs ffsss
 101 111 121 131 141
 GAAQMDGAIL VVAATDGPMP QTR~~EHILL~~SK CVGVPKMVVF LNKIDL~~EE~~EE
 bbbbbebbbb bbbbeeeee ebeebbbbbe ebeebbbb bbebeeeee
 sssssfs s s s ff ff fsffs ss fs s s sf f f ff
 151 161 171 181 191
 EEMVDLVEVE IRELLSSYDF DGDNTPIIRG SARGALEGKPE EAKVLEIM
 eebbebeee beebbeeeeb eeeebbbb bbbbebeee ebeeebeebb
 f ssf f f ss f ff f ss s
 201 211
 DAVDSYIDSP VREM
 ebbbebeee eeee
 f f f f

1243
1244
1245
1246

1247 **Sa_{EF-Tu} (Uniprot #: Q2G0N0)**1248 **Full length protein: 1-394**

| | | | | |
|------------------------------|---------------------|---------------------|---------------------|------------|
| 1 | 11 | 21 | 31 | 41 |
| MAK KE KFDRSK | EHANIGTIGH | VDHGKTTLTA | AIATVLAKNG | DSVAQSYDMI |
| 51 | 61 | 71 | 81 | 91 |
| DNAPEEKERG | ITINTSHIEY | Q TD KRHYAHV | DCPGHADYVK | NMITGAAQMD |
| 101 | 111 | 121 | 131 | 141 |
| GGILVVSAA D | GPMPQ TRE HI | LLSRNVGVPA | LVVFLNKVDM | VDDEELLELV |
| 151 | 161 | 171 | 181 | 191 |
| EMEV R DLLSE | YDFPGDDVPV | IAGSALKALE | GDAQYEE IL | ELMEAVDTYI |
| 201 | 211 | 221 | 231 | 241 |
| PT PER DS DK P | FMPV ED VFS | ITGRGTVA TG | RVERGOIKVG | EEVEIIGLHD |
| 251 | 261 | 271 | 281 | 291 |
| TSKTTV TG VE | MR K LLDYAE | AGDNIGALLR | GVARE D VORG | QVLAAPGSIT |
| 301 | 311 | 321 | 331 | 341 |
| PHTEFKAEVY | VLSKDEGR H | TPFFSNYRPQ | FYFRTTDV TG | VVHLPEGTEM |
| 351 | 361 | 371 | 381 | 391 |
| VMPGDNVEMT | VELI APIA IE | DGTRFSIREG | GRTVGS GV VT | EIIK |

1249
12501251 **Fragment 3:1-266**

1252

| | | | | |
|---------------------|---------------------|---------------------|--------------------|-------------|
| 1 | 11 | 21 | 31 | 41 |
| MA KE KFDRSK | EHANIGTIGH | VDHGKTTLTA | AIATVLAK NG | DSVAQSYDMI |
| eeeeeeeeee | eebbbbbb | bebebebe | bbbbeee | eebeeeeb |
| f f f f | f s sssss | ffsfsfsfs | ss | s |
| 51 | 61 | 71 | 81 | 91 |
| DNAPEEKERG | ITINTSHIEY | Q TD KRHYAHV | DCPGHADYVK | NMITGAAQMD |
| eebeeeee | bbbbbbbe | eeeebbbbb | bbbeebbe | ebbbbbbb |
| f sfff fs | sss f fs | f fssss | ssfffsfs | f fssssssss |
| 101 | 111 | 121 | 131 | 141 |
| GGILVVSAA D | GPMPQ TRE HI | LLSRNVGVPA | LVVFLNKVDM | VDDEELLELV |

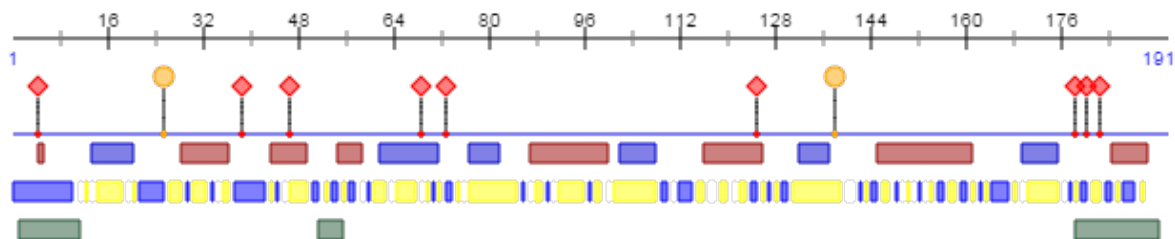
bbbbbbbbee eeeeebeebb bbbeebbeb bbbbbbebee eeeeebbee
 s s s s f f f fsffs ss ffs s s sf f fff f ss
 151 161 171 181 191
 EMEVRDILLSE YDFPGDDVPV IAGSALKALE GDAQYEEKIL ELMEAVDTYI
 eeebeebbee eebeeeebbeb bebbbbebbe eeeeeeebe ebbeebbee
 f f f ss f ff f ss s f

201 211 221 231 241
 PTPSRDSDKP FMMPVEDVFS ITGRGTVAATG RVERGQIKVG EEVEIIGHD
 eeeeeeeeee bbbebeebbbb bbbeebbeb ebbeebbeb ebbeebbbb
 f f f f s ff s sfsfs fs f f s f s

251 261
 TSKTTVIGVEMFQKLL
 eeeeebebbe bbeeee
 f f ssff f

1253
1254
1255
1256

Fragment 4: 1-192



1257
1258
1259 1 11 21 31 41
 MAKEKFDRSK EHANIGTIGH VDHGKTTLTA AIATVLA~~K~~NG DSVAQSYDMI
 eeeeeeeeee eebbbbbbb beebbebbeb bbbbeebbee eeebeebbeb
 f f f f f s sssss ffsfsfsfs ss s 51

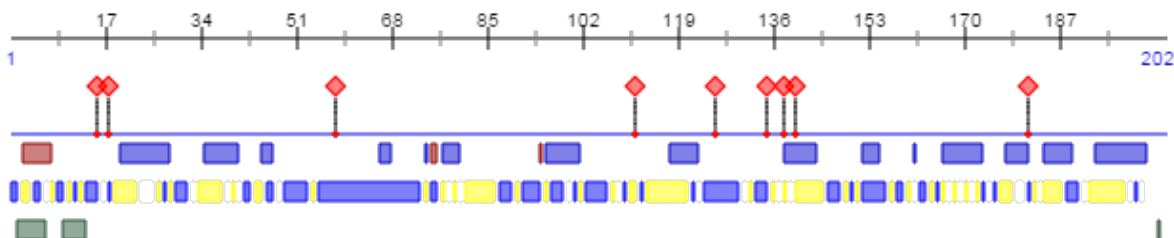
1262 61 71 81 91
 DNAPEEKERG ITINTSHIEY QTDKRYAHV DCPGHADYVK NMITGAAQMD
 eebeeeeee bbbbbbbeb eeeeebbbb bbeebbebbe ebbsbbbbb
 f sfff fs sss f fs f fssss s fffsfs f fsssssssssf 101

1266 111 121 131 141
 GGILVVSAAAD GPMPQTREHI LLSRNVG~~V~~PA LVVFLNKVDM VDDEELLELV
 bbbbbbbbee eeeeebeebb bbbeebbeb bbbbbbebee eeeeebbee
 s s s s f f fffsffs ss fsfs s sf f f f f ss 151

1270 161 171 181 192
 EMEVRDILLSE YDFPGDDVPV IAGSALKALE GDAQYEEKIL EL
 eeebeebbee eebeeeebbeb bebbbbebbe eeeeeeeeee ee
 f f f ss f ff f ss

1274

1275 **Fragment 5: 193-394**



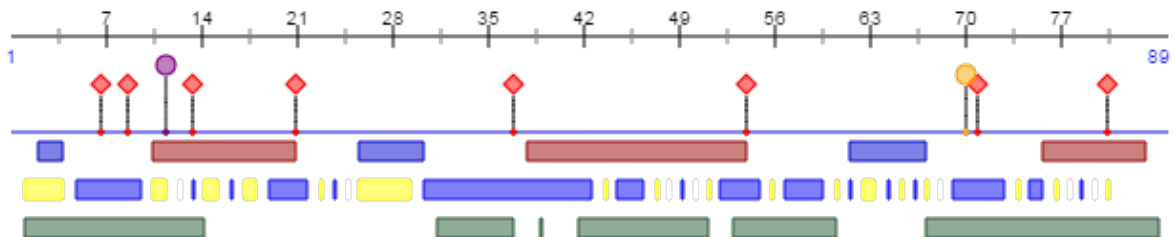
1276

| | | | | |
|-------------|------------|-------------|------------|------------|
| 193 | 203 | 213 | 223 | 233 |
| MEAVDTYIPT | PERDSKPFM | MPVEDVFSIT | GRGTVATGRV | ERGOIKVGEE |
| eeebbeebee | eeeeeeebb | bebeebbbb | bebebbebeb | eebebebee |
| f f f f | s | f ff s s | sfsfs fsf | f s |
| 243 | 253 | 263 | 273 | 283 |
| VEIIGLHDT | KTTVTGVEMF | RKLLDYAEAG | DNIGALLRGV | AREDVORGQV |
| bebbbbeeee | eeebbbb | eebbebeeee | eebebbb | eeeebeeb |
| s | f fss | ff f ff | ff f s fs | fs s |
| 293 | 303 | 313 | 323 | 333 |
| LAAPGSITFH | TEFKAEVYVL | SKDEGGRHTP | FFSNYPOFY | FRTTDVTGVV |
| beeeeebeeee | eebebebbb | eeeeeeeeeee | ebeeeeebb | bebeeebeb |
| f ff | s s | f fffffff | f ffffss | fsffffs |
| 343 | 353 | 363 | 373 | 383 |
| HLPEGTEMVM | PGDNVEMTVE | LIAPIAIEDG | TRFSIREGGR | TVGSGVTEI |
| eeebbbb | eeeebebebe | bbbbbbbeb | bebbb | bbbbbbbeb |
| s fs s | ffff | s | s sffff | sss s |
| 393 | | | | |
| IK | | | | |
| ee | | | | |

1277

1278

1279 **Fragment 6: 104-192**



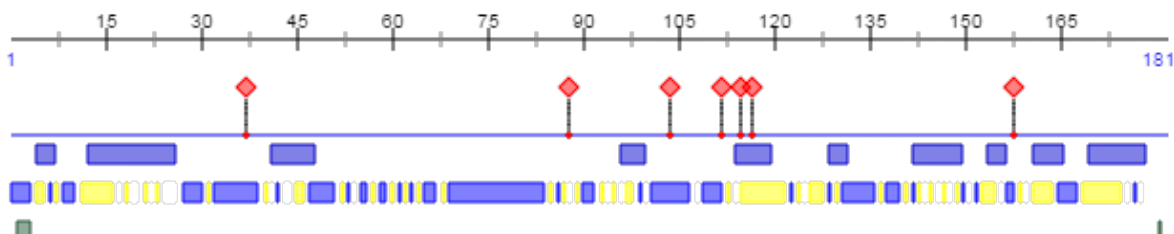
1280

| | | | | |
|------------|------------|-----------|------------|------------|
| 104 | 114 | 124 | 134 | 144 |
| LVVSAADGPM | POTREHILLS | RNVGVPAVV | FLNKVDMVDD | EELLELVEME |
| eebbb | eebeebbbb | eebeebbbb | bbbebeeee | eebebebee |
| ff s ff f | fs fssss | fsfs | s sf f f f | f sss f |

154 VRDLLSEYD 164 PGDDVPVIAG 174 SALKALEGDA 184 QYEELILEL
 beebbeeeeb eeeebbbbeb bbbbebeeee ebbeebbee
 f s ff f ss s f f

1281
1282

Fragment 7: 214-394



1283
1284
1285
1286
1287
1288
1289
1290
1291

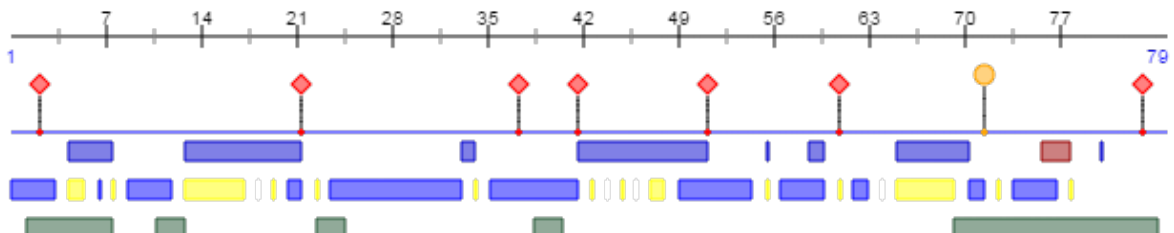
214 PVEDVFSITG 224 RGTVATGRVE 234 RGQIKVGEEV 244 EIIGLHDTSK 254 TTVTGVEMFR
 eeeebbbbb ebebbebe ebbebbbeb ebbbbeeeee eebebbebbe
 f ff s s s fsfs fsf f s s f fssf
 264 KLLDYAEAGD 274 NIGALLRGVA 284 REDVQRGQVL 294 AAPGSITPHT 304 EFKAEVVLS
 ebbeeeeee ebbbbbebe eeebeebbb eeeeebeeee ebbebbbbb
 f f fff f s s fs fs s f ff s s s 31

1292
1293
1294
1295
1296
1297
1298
1299

4 KDEGGRHTPF 324 FSNYPQFYF 334 RTTDVIGVVH 344 LPEGTEMVMP 354 GDNVEMTVEL
 eeeeeeeee beeeeebbb ebbebebe beebbebbe eebebebeb
 f ffffffff ffffs fsff fs s fs sf fff
 364 IAPIAIEDGT 374 RFSIREGGRT 384 VGSGVVTEII 394 K
 bbbbbbeeb ebbbeeeeb bbbbbbebe e
 s s fff s ss s

1300
1301
1302

Fragment 9: 214-292



1303

214 PVEDVFSITG 224 RGTVATGRVE 234 RGQIKVGEEV 244 EIIGLHDTSK 254 TTVTGVEMFR
 eeeebbbbb ebebbebe ebbebbbeb ebbbbeeeee eebebbebbe
 f ff s s s fsfs ssf f s f s fs fss

264 274 284
KLLDYAEAGD NIGALLRGVA REDVQRGV
ebbeeeeeee ebbbbebbe eeeeeeeee
f sf fff f s s fs ffff

1304

1305

1306

1307

1308

1309

1310 Will be presented as “Supplementary file S10- Experimental section”

1311 **S10. Supplementary Materials 3: Experimental Section**

1312 *S10.1. Host and human proteins used in binding assays*

1313 Host proteins used for affinity chromatography include: purified fibronectin (Code: 341635) and
1314 plasminogen (Code: 528175) from human plasma supplied by Merck Millipore; plasminogen
1315 from human plasma (Code: P7999) bovine actin (Code: A3653) and fetuin (Code: F3004)
1316 supplied by Sigma.

1317 Human proteins used for ELISA include: plasminogen (Code: P7999), lactoferrin (Code:
1318 L1294), laminin (Code: L6274), vitronectin (Code: SRP3186), plasma fibrinogen (Code: F3879)
1319 and plasma fibronectin (Code: 11051407001) supplied by Sigma.

1320 *S10.2. Peptide search parameters*

1321 Files were searched against the MSPnr100 database¹⁵¹ with the following parameters. Fixed
1322 modifications: none. Variable modifications: propionamide, oxidized methionine, deamidation.
1323 Enzyme: semi-trypsin. Number of allowed missed cleavages: 3. Peptide mass tolerance: 100
1324 ppm. MS/MS mass tolerance: 0.2 Da. Charge state: 2+, 3+ and 4+. For samples collected from
1325 the ‘Biotinylation enrichment of surface proteins’, ‘Avidin purification of A549 interacting
1326 proteins’ and ‘Avidin purification of PK-15 interacting proteins’, variable modifications also
1327 included NHS-LC-Biotin (K) and NHS-LC-Biotin (N-term). ‘Avidin purification of A549
1328 interacting proteins’ was also searched against *homo sapiens* entries in MSPnr100 to identify
1329 biotinylated surface A549 proteins. ‘Avidin purification of PK-15 interacting proteins’ was also
1330 searched against *sus scrofa* entries in MSPnr100 to identify biotinylated surface PK-15 proteins.

1331 *S. aureus* proteins were also searched against a *S. aureus* NCTC 8325 database derived from the
1332 published genome.

1333 ***S10.3. Expression and purification of rMpn_{Ef-Tu}***

1334 Expression and purification of rMpn_{Ef-Tu} was performed in one of two methods.

1335 The first method was performed as described in ¹⁰⁰. In brief, the *M. pneumoniae tuf* gene
1336 (*MPN_665*) was amplified and cloned with a N-terminal hexahistadine tail into the plasmid
1337 vector pET30 (Merck Millipore) containing a kanamycin resistance gene. The recombinant
1338 construct was transformed in to BL21-DE3 competent *E. coli* cells (Merck Millipore), induced
1339 with 1 mM isopropyl-β-D-thiogalactosidase (IPTG, Roth) and purified under denaturing
1340 conditions with immobilized metal affinity chromatography Ni²⁺-charged resin (Qiagen, Hilden,
1341 Germany) as described by the manufacturer. Elutions were concentrated using a 30 kDa
1342 Vivaspin centrifugal device (Sartorius, Göttingen, Germany). Recombinant protein was assayed
1343 and stored at -20°C. rMpn_{Ef-Tu} was used to produce guinea pig antiserum as reported ¹⁰⁰.

1344 The second method was performed as described in ⁸⁸ with modifications. In brief, the *M.*
1345 *pneumoniae tuf* gene (*MPN_665*) was synthesized with an N-terminal hexahistadine tail and
1346 cloned by Blue Heron Biotech (WA, USA) into a plasmid vector (PS100030) with an ampicillin
1347 resistance gene. The recombinant construct was transformed in to BL21-DE3 competent *E. coli*
1348 cells (Bioline, Eveleigh, Australia), induced with 1 mM isopropyl-β-D-thiogalactosidase (IPTG,
1349 Bioline) and purified under denaturing conditions with Profinity immobilized metal affinity
1350 chromatography Ni²⁺-charged resin (Bio-Rad, Gladesville, Australia). BL21 cells were lysed in 8
1351 M Urea, 100 mM Na₂HCO₃, 10 mM Tris-HCl, pH 8 with 6 rounds of sonication for 30 seconds
1352 on ice. Following centrifugation, the supernatant was added to Ni²⁺ resin overnight at 4°C. The

1353 resin was then loaded onto a column, washed four times with 5 ml 8 M Urea, 100 mM Na₂HCO₃,
1354 10 mM Tris-HCl, pH 6.3 and a 2-step elution: three times with 5 ml 8 M Urea, 100 mM
1355 Na₂HCO₃, 10 mM Tris-HCl, pH 5.9 and twice with 10 ml 8 M Urea, 100 mM Na₂HCO₃, 10 mM
1356 Tris-HCl, pH 4.5. Washes and elutions were monitored by SDS-PAGE. Elutions were
1357 concentrated using a 10 kDa Microsep™ centrifugal device (Pall, Port Washington, NY) and
1358 dialysed into PBS, 0.5% Tween 20 with 10,000 MWCO SnakeSkin® Dialysis Tubing (Thermo
1359 Fisher Scientific) at 4°C. Recombinant protein was assayed and stored at 4°C.

1360 *S10.4 LC-MS/MS of dimethyl labelled proteins*

1361 **S10.4.1. LC-MS/MS (Sciex 5600) of dimethyl labelled proteins**

1362 Peptides from dimethyl labelled proteins described in section 1.16.1 were separated by nanoLC
1363 using an Ultimate nanoRSLC UPLC and autosampler system (Dionex, Amsterdam,
1364 Netherlands). Samples (2.5 µl) were concentrated and desalted onto a micro C18 precolumn (300
1365 µm x 5 mm, Dionex) with H₂O:CH₃CN (98:2, 0.1 % TFA) at 15 µl/min. After a 4 min wash the
1366 pre-column was switched (Valco 10 port UPLC valve, Valco, Houston, TX) into line with a
1367 fritless nano column (75µ x ~15cm) containing C18AQ media (1.9µ, 120 Å Dr Maisch,
1368 Ammerbuch-Entringen Germany). Peptides were eluted using a linear gradient of H₂O:CH₃CN
1369 (98:2, 0.1 % formic acid) to H₂O:CH₃CN (64:36, 0.1 % formic acid) at 200 nl/min over 240 min.
1370 High voltage 2000 V was applied to low volume Titanium union (Valco) with the tip positioned
1371 ~ 0.5 cm from the curtain plate (T=150°C) of a 5600⁺ mass spectrometer (Sciex, Toronto,
1372 Canada). Positive ions were generated by electrospray and the 5600⁺ operated in information
1373 dependent acquisition mode (IDA).

1374 A survey scan m/z 350-1750 was acquired (PWHH resolution $\sim 30,000$, 0.25 sec acquisition
1375 time) with autocalibration enabled (at ~ 6 hr intervals). Up to the 10 most abundant ions (>300
1376 counts) with charge states $> +2$ and $< +5$ were sequentially isolated (width $m/z \sim 3$) and
1377 fragmented by CID with an optimal CE chosen based on m/z (product ion spectra were acquired
1378 at a resolution $\sim 20,000$ PWHH in 0.15 sec). M/z ratios selected for MS/MS were dynamically
1379 excluded for 30 or 45 seconds.

1380 Peak lists were generated using Mascot Daemon/Mascot Distiller (Matrix Science, London,
1381 England) or ProteinPilot (Sciex, v4.5) using default parameters, and submitted to the database
1382 search program Mascot (version 2.5.1, Matrix Science). Search parameters were: Precursor
1383 tolerance 10 ppm and product ion tolerances ± 0.05 Da; oxidation (M), deamidation (NQ),
1384 propionamide (C), Dimethyl (K), Dimethyl (N-term) specified as variable modifications; enzyme
1385 specificity was semi-ArgC; 1 missed cleavage was possible and the non-redundant protein
1386 database from NCBI (Jan 2015) searched.

1387 **S10.4.2. LC-MS/MS (Thermo Scientific Q Exactive™) of dimethyl labelled proteins**

1388 Peptides from dimethyl labelled proteins described in section 1.18.1 were separated by nanoLC
1389 using an Ultimate nanoRSLC UPLC and autosampler system (Dionex, Amsterdam,
1390 Netherlands). Samples (2.5 μ l) were concentrated and desalted onto a micro C18 precolumn (300
1391 μ m x 5 mm, Dionex) with $H_2O:CH_3CN$ (98:2, 0.1 % TFA) at 15 μ l/min. After a 4 min wash the
1392 pre-column was switched (Valco 10 port UPLC valve, Valco, Houston, TX) into line with a
1393 fritless nano column (75 μ x ~ 35 cm) containing C18AQ media (1.9 μ , 120 Å Dr Maisch,
1394 Ammerbuch-Entringen Germany). Peptides were eluted using a linear gradient of $H_2O:CH_3CN$
1395 (98:2, 0.1 % formic acid) to $H_2O:CH_3CN$ (64:36, 0.1 % formic acid) at 200 nl/min over 30 or

1396 240 min. High voltage 2000 V was applied to low volume Titanium union (Valco) with the
1397 column oven heated to 45°C (Sonation, Biberach, Germany) and the tip positioned ~ 0.5 cm
1398 from the heated capillary (T=300°C) of a QExactive Plus mass spectrometer (Thermo Fisher
1399 Scientific, Bremen, Germany). Positive ions were generated by electrospray and the QExactive
1400 operated in data dependent acquisition mode (DDA).

1401 A survey scan m/z 350-1750 was acquired (resolution = 70,000 at m/z 200, with an AGC target
1402 value of 10^6 ions) and lockmass was enabled (m/z 445.12003) Up to the 10 most abundant ions
1403 (>80,000 counts, underfill ratio 10%) with charge states > +2 and <+7 were sequentially isolated
1404 (width m/z 2.5) and fragmented by HCD (NCE = 30) with a AGC target of 10^5 ions (resolution =
1405 17,500 at m/z 200). M/z ratios selected for MS/MS were dynamically excluded for 30 or 45
1406 seconds.

1407 Peak lists were generated using Mascot Daemon/Mascot Distiller (Matrix Science, London,
1408 England) or Proteome Discoverer (Thermo Fisher Scientific, v1.4) using default parameters, and
1409 submitted to the database search program Mascot (version 2.5.1, Matrix Science). Search
1410 parameters were: Precursor tolerance 4 ppm and product ion tolerances ± 0.05 Da; oxidation
1411 (M), deamidation (NQ), propionamide (C), Dimethyl (K), Dimethyl (N-term) specified as
1412 variable modifications; enzyme specificity was semi-ArgC; 1 missed cleavage was possible and
1413 the non-redundant protein database from NCBI (Jan 2015) searched.

1414 **References**

- 1415 149 Yachdav, G. et al. PredictProtein--an open resource for online prediction of protein structural and
1416 functional features. *Nucleic acids research* 42, W337-343, doi:10.1093/nar/gku366 (2014).
- 1417 150 Hönigschmid, P. Improvement of DNA- and RNA protein binding prediction., Technische
1418 Universität München, (2012).
- 1419 151 Perkins, D. N., Pappin, D. J. C., Creasy, D. M. & Cottrell, J. S. Probability-based protein
1420 identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20, 3551-
1421 3567, doi:10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2 (1999).