Figure S1

A

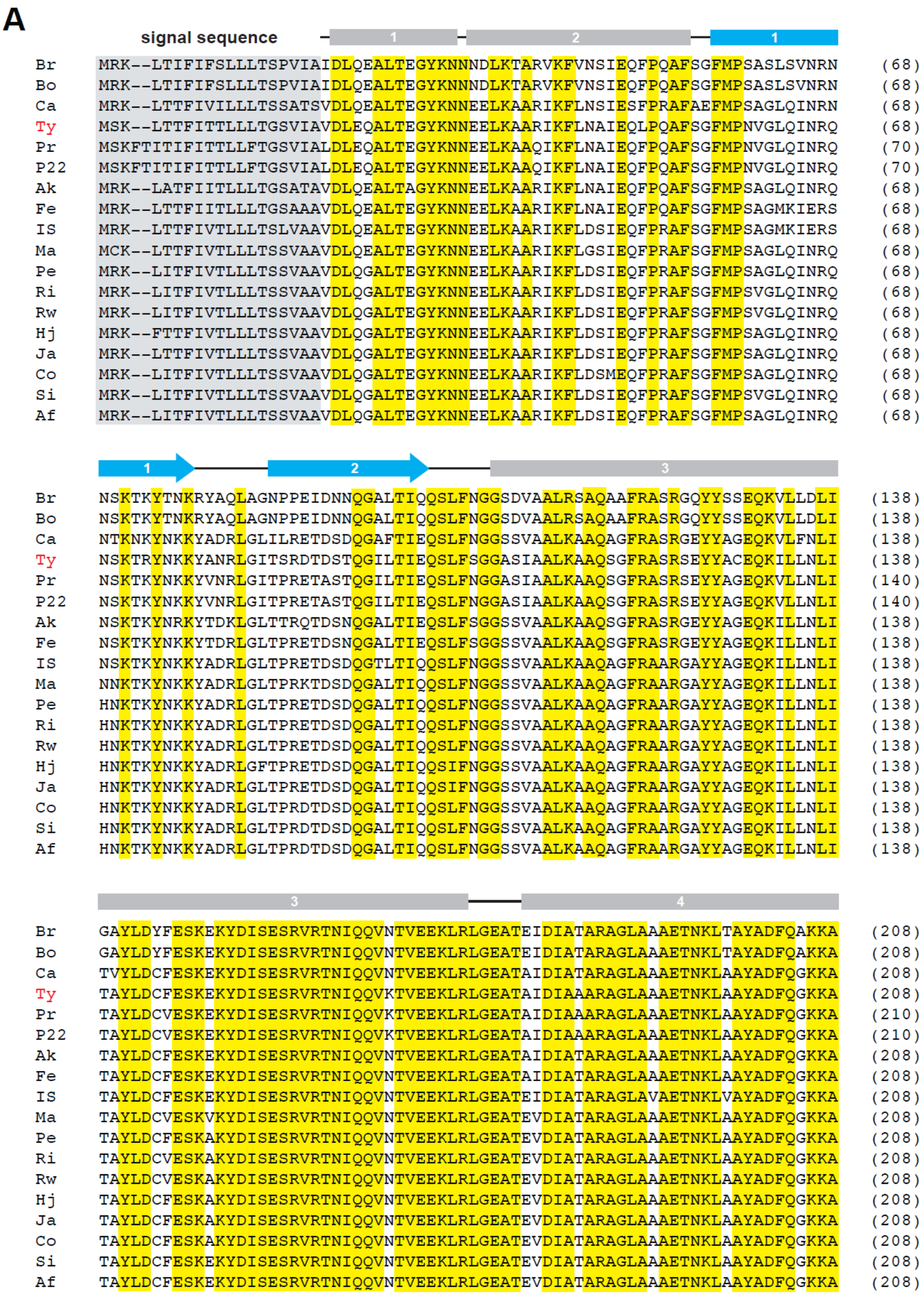
```
Rt      -WIASMGKRKLVESFIKYDI-DINKNKDI YSN-LSYYSAVDN-AFRAVARNNL------------FELRALLDVY-PILQEKNS
1       --LAYFGIDQLLERFLRGKL-DINA-KDN HGT-TALHWAVRN-GHKAMDKHGATPLHYSASNAHGTIVRVLLDKG-ANINKEER
2       ------GSTAIIEQLISSNV-DKDI-KNN RGE-TPLHIALQQHSSKD------------------KLIELLKALK-VNLQSKDS
3       -----------IGHLVDVGA-DKDI-TDS SGQ-TILQYALQKINRPN-------------------FQKLLSALG-ININEKNR
4       -FAAKNNSKESAELLISYGI-NINE-KND DGE-TAIHIAALF-NSKE-------------------TAELLISHG-ANINEKND
5       HIAAKNNRKEIVEFLISHGA-NINE-KDE DGK-TELHIAAEN-NSKA-------------------TAEVLISHG-ANINEKDE
6       --------KEIAEFLISHGA-NINE-KDE DGK-TELHIAAEN-NSKA-------------------TAEVLISHG-ANINEKDE
7       ---AAGGHMEVMEYLIQQGS-DLNK-NGN DGW-TPLHAAISN-GHLE-------------------VVNVLFAEG---AQGTRF
8       -YAAERGDINVVKLLLSYDI-DINE-KDV YGK-TPLHYAADN-IDKN-------------------IMELFTSKG-ANIEEKDL
9       LIASHFNHKDIVEFLFSHGA-DINI-KDN DGN-SPLHTT----SSVE-------------------LAEFLISHH-LDVNAKNI
10      -FAAENDNKEIAEILLLHGA-NINA-KDI YGN-TALHIAVEN-NSKE-------------------TAKILLLHG-ADINEKND
11      -YAAENGHRTVVKLFLARDIMNLDS-RDF TGS-TPFSRAAEN-GHVA-------------------VVRMLLKTGRVDVNSEDS
12      --VALKGRLNIVKEFFERKDFDHNA-RNK DGL-TILNAAIIS-GNDD-------------------LVYYIFDKY-PQLEWTDT
13      ---QREYLKKQIVFVIENKQFDQYY-DFL RGNVSEFAFAAFELGSSE-------------------ALKILYNIG-YPINIKNQ

Rt      TGETLLTAAIYNDNYYLAKFLVIRG-IKTSVLKKDE CKY-PLDIASSQGNTNIVCMLMKAKGYN---
1       SGRTALSRAIENGSQEIVKMLLDKK-ANVNRPKQST FIRTPLSLAAENGHGSVVKLLVE--------
2       NGYTLLHTAILEEDERLVSLLLNST-LAVDKNAKND FGKSPLHIAAEKGNLRLVNLLVALK------
3       NQQTLLHQSILEGNHELAKQLIAAG-ADIQ--AKNK QEYTPLHLAAIGGHLELVALLIAKDKAKNPN
4       NGETALHIAALNDNKETAELLISYG-IHIN--EKND NGETPLHIAEQYNNEEIAKLLI---------
5       YGQTALHIAAYNDSKEIAEFLISHG-ANIN--EKDE DGKTELHIAAENNSKATAEVLISHGAN----
6       YGQTALHIAAYNDSKEIAEFLISHG-ANIN--EKDE YGQTALHIATENNSKEIAEVLIE--------
7       EGLTLLYIASRYDHVDVVKFLVSKG-CDVN--EKSE CGKSPLHAACYTGNMNIVKLLVHHKA-----
8       NGKTPLHYAVPNRDIEITEFLFSNG-ANIN--DKDK NGKTPLLIASHFNHKDIVEFL----------
9       NGKTLLHFAATNNYKNMIEFLILHG-ANIN--EKDN DGKIPLHYAATANKANSECLITHGANIN---
10      NGQTALHIAVDNNTKKITEILLLHG-ANIN--EKDK MEKTPLQIATENDRKKIVKLLL---------
11      EGRTPLSLAAENGHVVVVKMLLKTGRLDID--LQDS EARTPLSWASGNGHATVVRALLK--------
12      LGQTPLHCAVANQNYVLTEYFLGKG-ANIN--TQDS TGRTPLHTACALGDIGLASILLNVQGIN---
13      DNETLLHAASKNNRTQLVTALLALG-IDVN--VRSN ENKSALYLAIEHGNVEITKLLLE--------
```

B

	Taxon	Acc. no.	Coords. (aa)	Bit score
Rt	*Rickettsia typhi* Wilmington	YP_067182	485-586	271
1	*Aspergillus flavus* NRRL3357	XP_002382730	180-314	41.6
2	Candidatus *Amoebophilus asiaticus* 5a2	YP_001957643	852-967	40
3	Candidatus *Amoebophilus asiaticus* 5a2	YP_001957643	1167-1280	41.6
4	*Trichomonas vaginalis* G3	XP_001579637	31-144	41.2
5	*Trichomonas vaginalis* G3	XP_001330150	149-268	41.2
6	*Trichomonas vaginalis* G3	XP_001330150	223-330	38.1
7	*Strongylocentrotus purpuratus*	XP_001196834	105-218	40.8
8	*Trichomonas vaginalis* G3	XP_001307918	309-421	40.4
9	*Trichomonas vaginalis* G3	XP_001307918	407-524	33.5
10	*Trichomonas vaginalis* G3	XP_001579859	417-530	40.4
11	*Penicillium marneffei* ATCC 18224	XP_002150094	190-307	40
12	*Trichomonas vaginalis* G3	XP_001308456	423-542	40
13	*Trichomonas vaginalis* G3	XP_001314500	1018-1133	37

Figure S1. **Structural prediction of the ankyrin repeats within RARP-1.** (**A**) Alignment of the Ankyrin (ANK) domain of *R. typhi* RARP-1 (NCBI accession no. YP_067182) with ANK domains of 13 diverse proteins retrieved from a PSI-BLAST search using only residues 485-586 as a query. ANK repeats 1, 2 and 3 are boxed black, tan and red, respectively. Structural model (6) for each repeat is mapped over the alignment. (**B**) Information for sequences shown in **A**. Coordinates of PSI-BLAST subjects producing significant alignments were excised and aligned with residues 485-586 of *R. typhi* RARP-1. Alignment constructed using MUSCLE v3.6 (3) with default parameters.

Figure S2

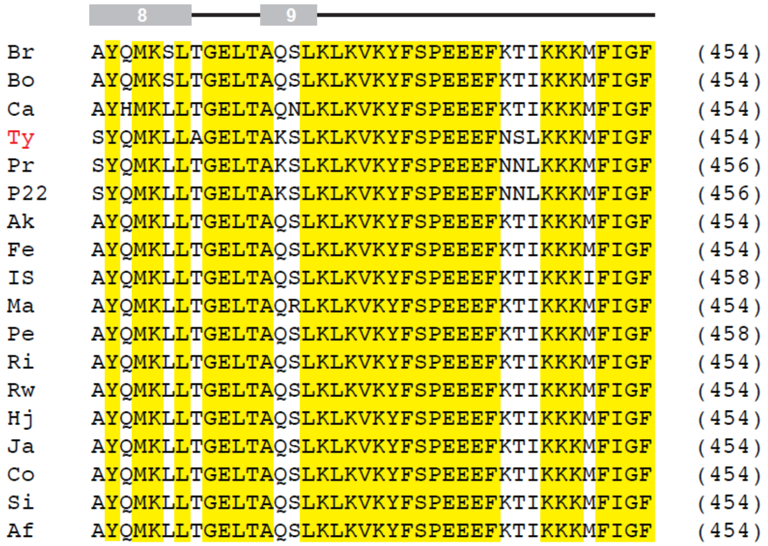


```
Br    MRK--LTIFIFSLLLTSPVIAIDLQEALTEGYKNNNDLKTARVKFVNSIEQFPQAFSGFMPSASLSVNRN   (68)
Bo    MRK--LTIFIFSLLLTSPVIAIDLQEALTEGYKNNNDLKTARVKFVNSIEQFPQAFSGFMPSASLSVNRN   (68)
Ca    MRK--LTTFIVILLLTSSATSVDLQEALTEGYKNNEELKAARIKFLNSIESFPRAFAEFMPSAGLQINRN   (68)
Ty    MSK--LTTFITTLLLTGSVIAVDLEQALTEGYKNNEELKAARIKFLNAIEQLPQAFSGFMPNVGLQINRQ   (68)
Pr    MSKFTITIFITTLLFTGSVIALDLEQALTEGYKNNEELKAAQIKFLNAIEQFPQAFSGFMPNVGLQINRQ   (70)
P22   MSKFTITIFITTLLFTGSVIALDLEQALTEGYKNNEELKAAQIKFLNAIEQFPQAFSGFMPNVGLQINRQ   (70)
Ak    MRK--LATFIITLLLTGSATAVDLQEALTAGYKNNEELKAARIKFLNAIEQFPQAFSGFMPSAGLQINRQ   (68)
Fe    MRK--LTTFIITLLLTGSAAAVDLQEALTEGYKNNEELKAARIKFLNAIEQFPQAFSGFMPSAGMKIERS   (68)
IS    MRK--LTTFIVTLLLTSLVAAVDLQEALTEGYKNNEELKAARIKFLDSIEQFPRAFSGFMPSAGMKIERS   (68)
Ma    MCK--LTTFIVTLLLTSSVAAVDLQEALTEGYKNNEELKAARIKFLGSIEQFPRAFSGFMPSAGLQINRQ   (68)
Pe    MRK--LITFIVTLLLTSSVAAVDLQGALTEGYKNNEELKAARIKFLDSIEQFPRAFSGFMPSAGLQINRQ   (68)
Ri    MRK--LITFIVTLLLTSSVAAVDLQGALTEGYKNNEELKAARIKFLDSIEQFPRAFSGFMPSVGLQINRQ   (68)
Rw    MRK--LITFIVTLLLTSSVAAVDLQGALTEGYKNNEELKAARIKFLDSIEQFPRAFSGFMPSVGLQINRQ   (68)
Hj    MRK--FTTFIVTLLLTSSVAAVDLQGALTEGYKNNEELKAARIKFLDSIEQFPRAFSGFMPSAGLQINRQ   (68)
Ja    MRK--LTTFIVTLLLTSSVAAVDLQGALTEGYKNNEELKAARIKFLDSIEQFPRAFSGFMPSAGLQINRQ   (68)
Co    MRK--LITFIVTLLLTSSVAAVDLQGALTEGYKNNEELKAARIKFLDSMEQFPRAFSGFMPSAGLQINRQ   (68)
Si    MRK--LITFIVTLLLTSSVAAVDLQGALTEGYKNNEELKAARIKFLDSIEQFPRAFSGFMPSVGLQINRQ   (68)
Af    MRK--LITFIVTLLLTSSVAAVDLQGALTEGYKNNEELKAARIKFLDSIEQFPRAFSGFMPSAGLQINRQ   (68)

Br    NSKTKYTNKRYAQLAGNPPEIDNNQGALTIQQSLFNGGSDVAALRSAQAAFRASRGQYYSSEQKVLLDLI  (138)
Bo    NSKTKYTNKRYAQLAGNPPEIDNNQGALTIQQSLFNGGSDVAALRSAQAAFRASRGQYYSSEQKVLLDLI  (138)
Ca    NTKNKYNKKYADRLGLILRETDSDQGAFTIEQSLFNGGSSVAALKAAQAGFRASRGEYYAGEQKVLFNLI  (138)
Ty    NSKTRYNKKYANRLGITSRDTDSTQGILTIEQSLFSGGASIAALKAAQSGFRASRSEYYACEQKILLNLI  (138)
Pr    NSKTKYNKKYVNRLGITPRETASTQGILTIEQSLFNGGASIAALKAAQSGFRASRSEYYAGEQKVLLNLI  (140)
P22   NSKTKYNKKYVNRLGITPRETASTQGILTIEQSLFNGGASIAALKAAQSGFRASRSEYYAGEQKVLLNLI  (140)
Ak    NSKTKYNRKYTDKLGLTTRQTDSNQGALTIEQSLFSGGSSVAALKAAQSGFRASRGEYYAGEQKILLNLI  (138)
Fe    NSKTKYNKKYTDRLGLTPRETDSNQGALTIEQSLFSGGSSVAALKAAQSGFRASRGEYYAGEQKILLNLI  (138)
IS    NSKTKYNKKYADRLGLTPRETDSDQGTLTIQQSLFNGGSSVAALKAAQAGFRAARGAYYAGEQKILLNLI  (138)
Ma    NNKTKYNKKYADRLGLTPRKTDSDQGALTIQQSLFNGGSSVAALKAAQAGFRAARGAYYAGEQKILLNLI  (138)
Pe    HNKTKYNKKYADRLGLTPRETDSDQGALTIQQSLFNGGSSVAALKAAQAGFRAARGAYYAGEQKILLNLI  (138)
Ri    HNKTKYNKKYADRLGLTPRETDSDQGALTIQQSLFNGGSSVAALKAAQAGFRAARGAYYAGEQKILLNLI  (138)
Rw    HNKTKYNKKYADRLGLTPRETDSDQGALTIQQSLFNGGSSVAALKAAQAGFRAARGAYYAGEQKILLNLI  (138)
Hj    HNKTKYNKKYADRLGFTPRETDSDQGALTIQQSIFNGGSSVAALKAAQAGFRAARGAYYAGEQKILLNLI  (138)
Ja    HNKTKYNKKYADRLGLTPRETDSDQGALTIQQSIFNGGSSVAALKAAQAGFRAARGAYYAGEQKILLNLI  (138)
Co    HNKTKYNKKYADRLGLTPRDTDSDQGALTIQQSLFNGGSSVAALKAAQAGFRAARGAYYAGEQKILLNLI  (138)
Si    HNKTKYNKKYADRLGLTPRDTDSDQGALTIQQSLFNGGSSVAALKAAQAGFRAARGAYYAGEQKILLNLI  (138)
Af    HNKTKYNKKYADRLGLTPRDTDSDQGALTIQQSLFNGGSSVAALKAAQAGFRAARGAYYAGEQKILLNLI  (138)

Br    GAYLDYFESKEKYDISESRVRTNIQQVNTVEEKLRLGEATEIDIATARAGLAAAETNKLTAYADFQAKKA  (208)
Bo    GAYLDYFESKEKYDISESRVRTNIQQVNTVEEKLRLGEATEIDIATARAGLAAAETNKLTAYADFQAKKA  (208)
Ca    TVYLDCFESKEKYDISESRVRTNIQQVNTVEEKLRLGEATAIDIATARAGLAAAETNKLAAYADFQGKKA  (208)
Ty    TAYLDCFESKEKYDISESRVRTNIQQVKTVEEKLRLGEATAIDIAAARAGLAAAETNKLAAYADFQGKKA  (208)
Pr    TAYLDCVESKEKYDISESRVRTNIQQVKTVEEKLRLGEATAIDIAAARAGLAAAETNKLAAYADFQGKKA  (210)
P22   TAYLDCVESKEKYDISESRVRTNIQQVKTVEEKLRLGEATAIDIAAARAGLAAAETNKLAAYADFQGKKA  (210)
Ak    TAYLDCFESKEKYDISESRVRTNIQQVNTVEEKLRLGEATAIDIATARAGLAAAETNKLAAYADFQGKKA  (208)
Fe    TAYLDCFESKEKYDISESRVRTNIQQVNTVEEKLRLGEATAIDIATARAGLAAAETNKLAAYADFQGKKA  (208)
IS    TAYLDCFESKEKYDISESRVRTNIQQVNTVEEKLRLGEATEIDIATARAGLAVAETNKLVAYADFQGKKA  (208)
Ma    TAYLDCVESKVKYDISESRVRTNIQQVNTVEEKLRLGEATEVDIATARAGLAAAETNKLAAYADFQGKKA  (208)
Pe    TAYLDCFESKAKYDISESRVRTNIQQVNTVEEKLRLGEATEVDIATARAGLAAAETNKLAAYADFQGKKA  (208)
Ri    TAYLDCVESKAKYDISESRVRTNIQQVNTVEEKLRLGEATEVDIATARAGLAAAETNKLAAYADFQGKKA  (208)
Rw    TAYLDCVESKAKYDISESRVRTNIQQVNTVEEKLRLGEATEVDIATARAGLAAAETNKLAAYADFQGKKA  (208)
Hj    TAYLDCFESKAKYDISESRVRTNIQQVNTVEEKLRLGEATEVDIATARAGLAAAETNKLAAYADFQGKKA  (208)
Ja    TAYLDCFESKAKYDISESRVRTNIQQVNTVEEKLRLGEATEVDIATARAGLAAAETNKLAAYADFQGKKA  (208)
Co    TAYLDCFESKAKYDISESRVRTNIQQVNTVEEKLRLGEATEVDIATARAGLAAAETNKLAAYADFQGKKA  (208)
Si    TAYLDCFESKEKYDISESRVRTNIQQVNTVEEKLRLGEATEVDIATARAGLAAAETNKLAAYADFQGKKA  (208)
Af    TAYLDCFESKEKYDISESRVRTNIQQVNTVEEKLRLGEATEVDIATARAGLAAAETNKLAAYADFQGKKA  (208)
```

```
Br    NFIRVFGIEPTDIAMPILPQGLPNSLDELTKRAVNLNPEIDSAKHSVTSAKAQELVEKGRLLPQVSVQLQ  (278)
Bo    NFIRVFGIEPTDIAMPILPQGLPNSLDELTKRAVNLNPEIDSAKHSVTSAKAQELVEKGRLLPQVSVQLQ  (278)
Ca    NFIRVFGIAPTNITMPDLPKMLPASLDELTKKAAKLNPDIDSARHNVTSAKASEMAAKGKLLPQVSVKLQ  (278)
Ty    NFIKVFGIEANDITMPDLPDRLPTSLDEFTRKAAKLNPDINSARHNVSVAKALEMVQKGKLLPQVSVKLL  (278)
Pr    NFIKVFGIEANDITMPDLPDRLPISLDEFTRKAAKFNPDINSARHNVTVTKALEMVQKGKLLPQVSVKLL  (280)
P22   NFIKVFGIEANDITMPDLPDRLPISLDEFTRKAAKFNPDINSARHNVTVTKALEMVQKGKLLPQVSVKLL  (280)
Ak    NFIRVFGIEATDLTMPNLPDKLPASLDELTRKAAKFNPDIDSARHNVISAKALEMVQKGKLLPQVSVKLQ  (278)
Fe    NFIRVFGIEATDITMPNLPDRLPASLDELTKKAAKLNPDIDSARHNVTSAKALEMVQKGKLLPQVSVKLQ  (278)
IS    NFIRVFGIEPTNITMPELPKMLPASLDELTKRAAKLNPDIDSARHNITSAKALEMAEKGKLLPQVSVQLQ  (278)
Ma    NFIRVFGIEPTNITMPELPKMLPASLDELTRRAAKLNPDIDSARHNVTSAKALEMAEKGKLLPQVSVQLQ  (278)
Pe    NFIKVFGIEPTNITMPELPKMLPASLDELTGRAAKLNPDIDSARHNVTAAKALEMAEKGKLLPQVSVQLQ  (278)
Ri    NFIRVFGIEPTNITMPELPKMLPASLDELTGRAAKLNPDIDSARHNVTAAKALEMAEKGKLLPQVSVQLQ  (278)
Rw    NFIRVFGIEPTNITMPELPKMLPASLDELTGRAAKLNPDIDSARHNVTAAKALEMAEKGKLLPQVSVQLQ  (278)
Hj    NFIRVFGIEPTNITMPELPKMLPASLDELTGRAAKLNPDIDSARHNVTAAKALEMAEKGKLLPQVSVQLQ  (278)
Ja    NFIRVFGIEPTNITMPELPKMLPASLDELTGRAAKLNPDIDSARHNVTAAKALEMAEKGKLLPQVSVQLQ  (278)
Co    NFIRVFGIEPTNITMPELPKMLPASLDELTRRAAKLNPDIDSARHNVTAAKALEMAEKGKLLPQVSVQLQ  (278)
Si    NFIRVFGIEPTNITMPELPKMLPASLDELTGRAAKLNPDIDSARHNVTAAKALEMAEKGKLLPQVSVQLQ  (278)
Af    NFIRVFGIEPTNITMPELPKMLPASLDELTGRAAKLNPDIDSARHNVTAAKALEMAEKGKLLPQVSVQLQ  (278)
```

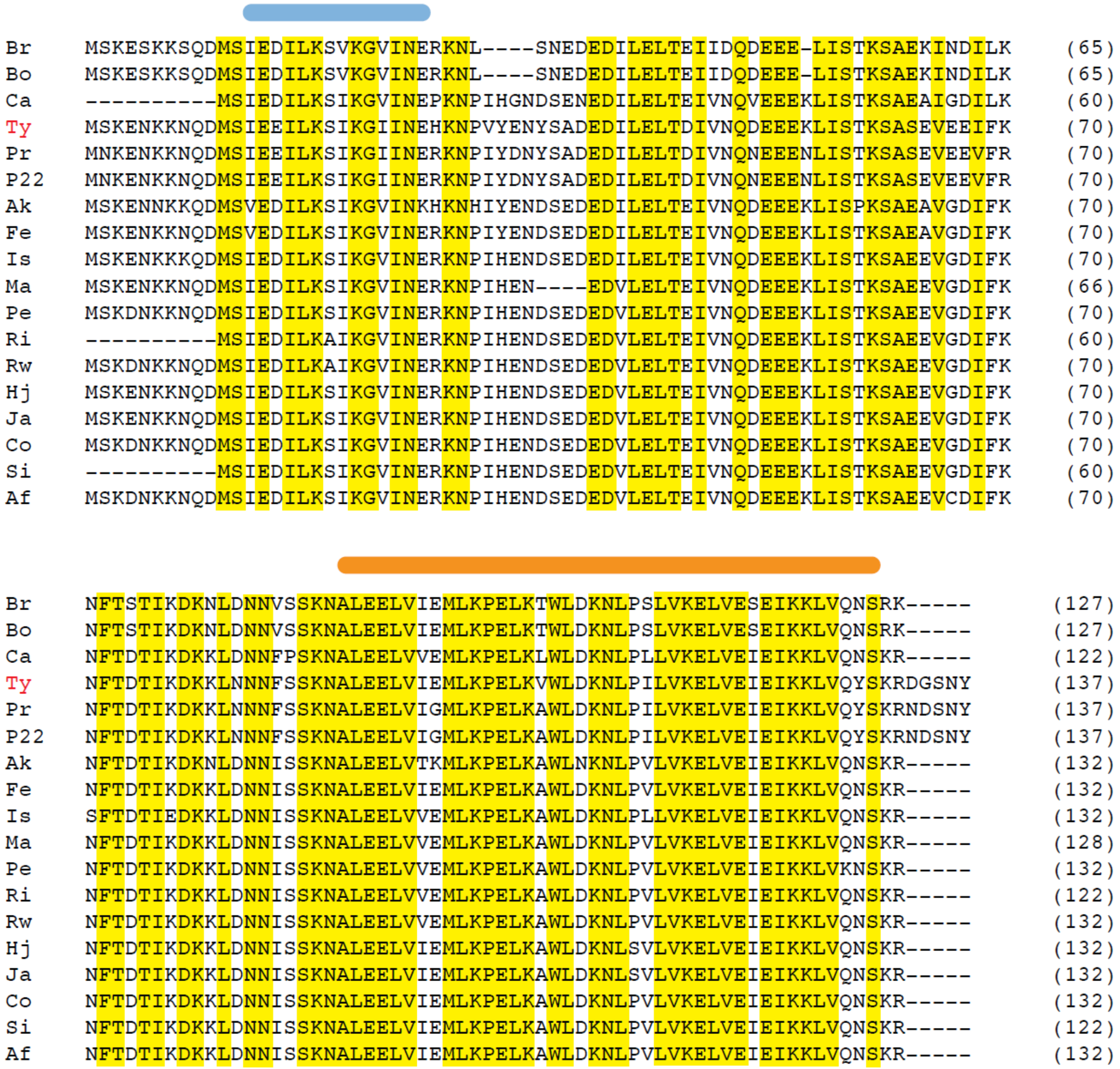
```
Br    SGKTNYNPQNLN----TQQINTRSVTTTLSVNVPIYPNGGAQYSKIRSAKNQTRNSAIQLDSVIRQTQAY  (344)
Bo    SGKTNYNPQNLN----TKQINTRSVTTTLSVNVPIYPNGGAQYSKIRSAKNQTRNSAIQLDSVIRQTQAY  (344)
Ca    SGRTHYNPQDPV----VQNINTKSVTTTLSVNIPIYPEGGAQYSRIRSAQNQTRNSAVQLDSAIKQTQAW  (344)
Ty    SGRTNYNPQEPV----IQNINNRIYTTTLAVNIPIYPEGGAQYSRIRSAKNQTRNSVVQLDSAIKQIKAG  (344)
Pr    SGGTNYNPQEPV----IQNINNRIYTTTLSVNIPIYPEGGAQYSRIRSAKNQTRNSVVQLDSAIKQIKAG  (346)
P22   SGGTNYNPQEPV----IQNINNRIYTTTLSVNIPIYPEGGAQYSRIRSAKNQTRNSVVQLDSAIKQIKAG  (346)
Ak    SGRTDYNPQDPA----VQNINNKSYTTILSVNIPIYPDGGAQYSRIRSAKNQTRSSAVQLDSAIKQTQAG  (344)
Fe    SGRTNYNPQDPA----VQNINNKSYTTILSVNIPIYPNGGAQYSEIRKAKNQTRNSAVQLDSAIKQTQAG  (344)
IS    SGRTHYNPQDDINARNVNAINNGSYTTTLSVNIPIYPEGGAQYSRIRSAKNQTRNSAVQLDSTIKQTQAG  (348)
Ma    SGRTHYNPQDDI-----NAINNRSYTTTLSVNIPIYPEGGAQYSRIRSAKNQTRNSAVQLDSAIKQTQAG  (344)
Pe    SGRTYYNPQGDINARNVNAINNRSYTTTLSVNIPIYPEGGAQYSRIRSAKNQTRNSAVQLDNVIKQTQAG  (348)
Ri    SGRTYYNPQGDI-----NAINNRSYTTTLSVNIPIYPEGGAQYSRIRSAKNQTRNSAVQLDNVIKQTQAG  (344)
Rw    SGRTYYNPQGDI-----NAINNRSYTTTLSVNIPIYPEGGAQYSRIRSAKNQTRNSAVQLDNVIKQTQAG  (344)
Hj    SGRTYYNPQDDINARNVNAINNRSYTTTLSVNIPIYPEGGAQYSRIRSAKNQTRNSAVQLDSVIKQTQAG  (348)
Ja    SGRTYYNPQDDINARNVNAINNRSYTTTLSVNIPIYPEGGAQYSRIRSAKNQTRNSAVQLDSVIKQTQAW  (348)
Co    SGRTYYNPQGDI-----NAINNRSYTTTLSVNIPIYPEGGAQYSRIRSAKNQTRNSAVQLDNVIKQTQAG  (344)
Si    SGRTYYNPQGDI-----NAINNRSYTTTLSVNIPIYPEGGAQYSRIRSAKNQTRNSAVQLDNVIKQTQAG  (344)
Af    SGRTYYNPQGDI-----NAINNRSYTTTLSVNIPIYPEGGAQYSRIRSAKNQTRNSAVQLDNVIKQTQAG  (344)
```

```
Br    VISIWEGFEAAKSRIIAADQGVAAAQISYDGTVQEEIVGSKTMLDVLSAEEKLYDAKITRVDAYKASILS  (414)
Bo    VISIWEGFEAAKSRIIAADQGVAAAQISYDGTVQEEIVGSKTMLDVLSAEEKLYDAKITRVDAYKASILS  (414)
Ca    VVSVWEGFEAAKSRIVAANQGVDAAQISYDGTVQEEIVGSKTILDVLTAEEKLYEAKITRVDAYKSLVLA  (414)
Ty    VVSVWEGFETAKSRIVAANQGVEAAQISYNGIVQEEIVGSKTILDVLDAEQKLYEAKITRVDAYKNSVLA  (414)
Pr    VVSVWEGFETAKSRIVAANQGVEAAQISYNGIVQEEIVGSKTILDVLDAEQKLYEAKITRVDAYKNSVLA  (416)
P22   VVSVWEGFETAKSRIVAANQGVEAAQISYNGIVQEEIVGSKTILDVLDAEQKLYEAKITRVDAYKNSVLA  (416)
Ak    VVSVWEGFEAAKSRIVAANQGVDAAQISYDGTVQEEIVGSKTILDVLDAEQKLYEAKITRVDAYKNSVLA  (414)
Fe    VVSVWEGFEAAKSRIVAANQGVEAAQISYDGTVQEEIVGSKTILDVLDAEQKLYEAKITRVDAYKNSVLA  (414)
IS    VVSVWEGFEAAKSRIVAANQGVDAAQISYDGTVQEEIVGSKTILDVLTAEEKLYEAKITRVDAYKSLVLA  (418)
Ma    VVSVWEGFEAAKSRIVAANQGVDAAQISYDGTVQEEIVGSKTILDVLTAEEKLYEAKITRVDAYKSSVLA  (414)
Pe    VVSVWEGFEAAKSRIVAANQGVDAAQISYDGTVQEEIVGSKTILDVLTAEEKLYEAKITRVDAYKSSVLA  (418)
Ri    VVSVWEGFEAAKSRIVAANQGVDAAQISYDGTVQEEIVGSKTILDVLTAEEKLYEAKITRVDAYKSSVLA  (414)
Rw    VVSVWEGFEAAKSRIVAANQGVDAAQISYDGTVQEEIVGSKTILDVLTAEEKLYEAKITRVDAYKSSVLA  (414)
Hj    VVSVWEGFEAAKSRIVAANQGVDAAQISYDGTVQEEIVGSKTILDVLTAEEKLYEAKITRVDAYKSSVLA  (414)
Ja    VVSVWEGFEAAKSRIVAANQGVDAAQISYDGTVQEEIVGSKTILDVLTAEEKLYEAKITRVDAYKSSVLA  (414)
```

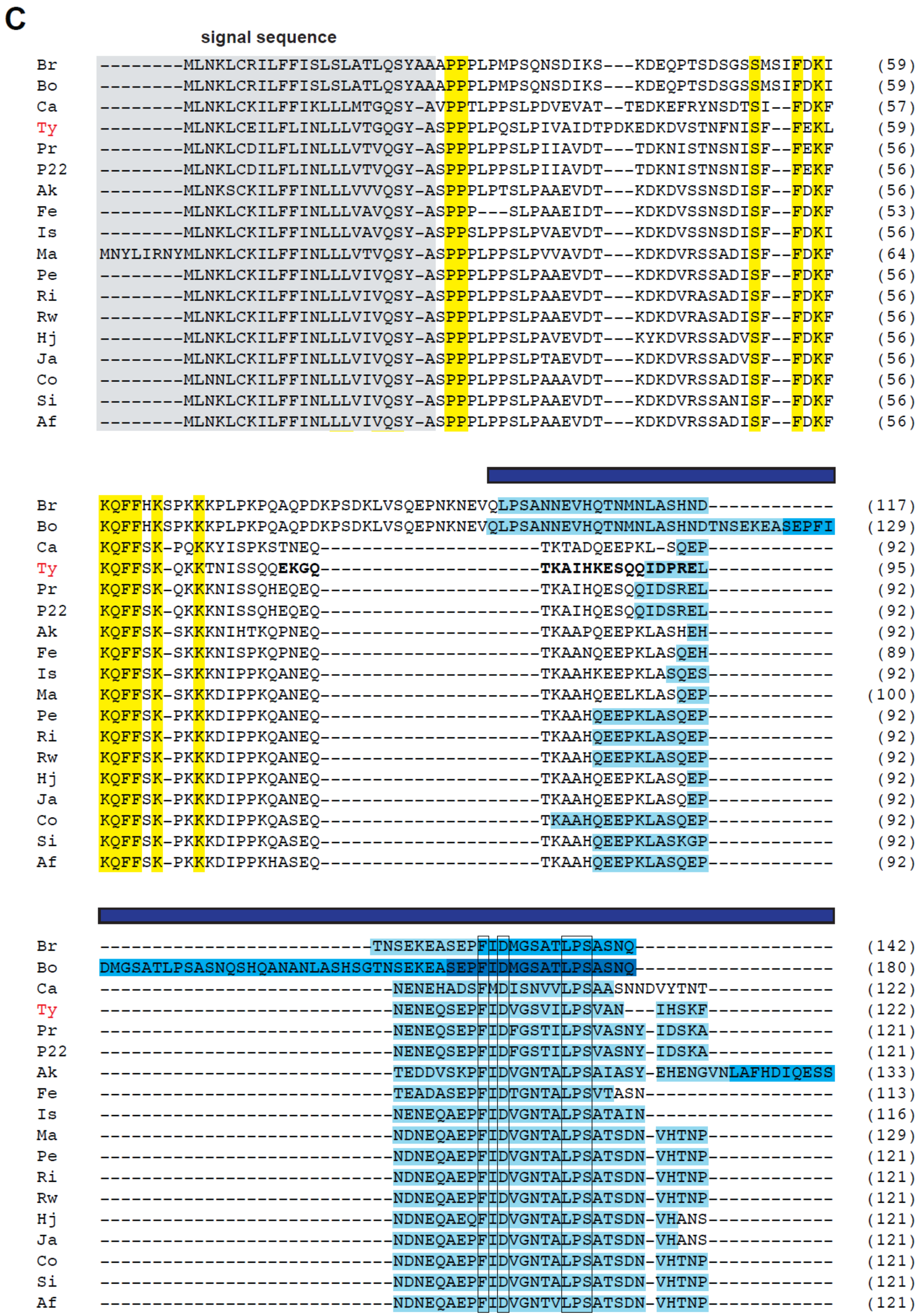


```
Br     AYQMKSLTGELTAQSLKLKVKYFSPEEEFKTIKKKMFIGF   (454)
Bo     AYQMKSLTGELTAQSLKLKVKYFSPEEEFKTIKKKMFIGF   (454)
Ca     AYHMKLLTGELTAQNLKLKVKYFSPEEEFKTIKKKMFIGF   (454)
Ty     SYQMKLLAGELTAKSLKLKVKYFSPEEEFNSLKKKMFIGF   (454)
Pr     SYQMKLLTGELTAKSLKLKVKYFSPEEEFNNLKKKMFIGF   (456)
P22    SYQMKLLTGELTAKSLKLKVKYFSPEEEFNNLKKKMFIGF   (456)
Ak     AYQMKLLTGELTAQSLKLKVKYFSPEEEFKTIKKKMFIGF   (454)
Fe     AYQMKLLTGELTAQSLKLKVKYFSPEEEFKTIKKKMFIGF   (454)
IS     AYQMKLLTGELTAQSLKLKVKYFSPEEEFKTIKKKIFIGF   (458)
Ma     AYQMKLLTGELTAQRLKLKVKYFSPEEEFKTIKKKMFIGF   (454)
Pe     AYQMKLLTGELTAQSLKLKVKYFSPEEEFKTIKKKMFIGF   (458)
Ri     AYQMKLLTGELTAQSLKLKVKYFSPEEEFKTIKKKMFIGF   (454)
Rw     AYQMKLLTGELTAQSLKLKVKYFSPEEEFKTIKKKMFIGF   (454)
Hj     AYQMKLLTGELTAQSLKLKVKYFSPEEEFKTIKKKMFIGF   (454)
Ja     AYQMKLLTGELTAQSLKLKVKYFSPEEEFKTIKKKMFIGF   (454)
Co     AYQMKLLTGELTAQSLKLKVKYFSPEEEFKTIKKKMFIGF   (454)
Si     AYQMKLLTGELTAQSLKLKVKYFSPEEEFKTIKKKMFIGF   (454)
Af     AYQMKLLTGELTAQSLKLKVKYFSPEEEFKTIKKKMFIGF   (454)
```

B



```
Br    MSKESKKSQDMSIEDILKSVKGVINERKNL----SNEDEDILELTEIIDQDEEE-LISTKSAEKINDILK    (65)
Bo    MSKESKKSQDMSIEDILKSVKGVINERKNL----SNEDEDILELTEIIDQDEEE-LISTKSAEKINDILK    (65)
Ca    ----------MSIEDILKSIKGVINEPKNPIHGNDSENEDILELTEIVNQVEEEKLISTKSAEAIGDILK    (60)
Ty    MSKENKKNQDMSIEEILKSIKGIINEHKNPVYENYSADEDILELTDIVNQDEEEKLISTKSASEVEEIFK    (70)
Pr    MNKENKKNQDMSIEEILKSIKGIINERKNPIYDNYSADEDILELTDIVNQNEEENLISTKSASEVEEVFR    (70)
P22   MNKENKKNQDMSIEEILKSIKGIINERKNPIYDNYSADEDILELTDIVNQNEEENLISTKSASEVEEVFR    (70)
Ak    MSKENNKKQDMSVEDILKSIKGVINKHKNHIYENDSEDEDILELTEIVNQDEEEKLISPKSAEAVGDIFK    (70)
Fe    MSKENKKNQDMSVEDILKSIKGVINERKNPIYENDSEDEDILELTEIVNQDEEEKLISTKSAEAVGDIFK    (70)
Is    MSKENKKKQDMSIEDILKSIKGVINERKNPIHENDSEDEDILELTEIVNQDEEEKLISTKSAEEVGDIFK    (70)
Ma    MSKENKKNQDMSIEDILKSIKGVINERKNPIHEN----EDVLELTEIVNQDEEEKLISTKSAEEVGDIFK    (66)
Pe    MSKDNKKNQDMSIEDILKSIKGVINERKNPIHENDSEDEDVLELTEIVNQDEEEKLISTKSAEEVGDIFK    (70)
Ri    ----------MSIEDILKAIKGVINERKNPIHENDSEDEDVLELTEIVNQDEEEKLISTKSAEEVGDIFK    (60)
Rw    MSKDNKKNQDMSIEDILKAIKGVINERKNPIHENDSEDEDVLELTEIVNQDEEEKLISTKSAEEVGDIFK    (70)
Hj    MSKENKKNQDMSIEDILKSIKGVINERKNPIHENDSEDEDVLELTEIVNQDEEEKLISTKSAEEVGDIFK    (70)
Ja    MSKENKKNQDMSIEDILKSIKGVINERKNPIHENDSEDEDVLELTEIVNQDEEEKLISTKSAEEVGDIFK    (70)
Co    MSKDNKKNQDMSIEDILKSIKGVINERKNPIHENDSEDEDVLELTEIVNQDEEEKLISTKSAEEVGDIFK    (70)
Si    ----------MSIEDILKSIKGVINERKNPIHENDSEDEDVLELTEIVNQDEEEKLISTKSAEEVGDIFK    (60)
Af    MSKDNKKNQDMSIEDILKSIKGVINERKNPIHENDSEDEDVLELTEIVNQDEEEKLISTKSAEEVCDIFK    (70)

Br    NFTSTIKDKNLDNNVSSKNALEELVIEMLKPELKTWLDKNLPSLVKELVESEIKKLVQNSRK-----    (127)
Bo    NFTSTIKDKNLDNNVSSKNALEELVIEMLKPELKTWLDKNLPSLVKELVESEIKKLVQNSRK-----    (127)
Ca    NFTDTIKDKKLDNNFPSKNALEELVVEMLKPELKLWLDKNLPLLVKELVEIEIKKLVQNSKR-----    (122)
Ty    NFTDTIKDKKLNNNFSSKNALEELVIEMLKPELKVWLDKNLPILVKELVEIEIKKLVQYSKRDGSNY    (137)
Pr    NFTDTIKDKKLNNNFSSKNALEELVIGMLKPELKAWLDKNLPILVKELVEIEIKKLVQYSKRNDSNY    (137)
P22   NFTDTIKDKKLNNNFSSKNALEELVIGMLKPELKAWLDKNLPILVKELVEIEIKKLVQYSKRNDSNY    (137)
Ak    NFTDTIKDKNLDNNISSKNALEELVTKMLKPELKAWLNKNLPVLVKELVEIEIKKLVQNSKR-----    (132)
Fe    NFTDTIKDKKLDNNISSKNALEELVIEMLKPELKAWLDKNLPVLVKELVEIEIKKLVQNSKR-----    (132)
Is    SFTDTIEDKKLDNNISSKNALEELVVEMLKPELKAWLDKNLPLLVKELVEIEIKKLVQNSKR-----    (132)
Ma    NFTDTIKDKKLDNNISSKNALEELVVEMLKPELKAWLDKNLPVLVKELVEIEIKKLVQNSKR-----    (128)
Pe    NFTDTIKDKKLDNNISSKNALEELVVEMLKPELKAWLDKNLPVLVKELVEIEIKKLVKNSKR-----    (132)
Ri    NFTDTIKDKKLDNNISSKNALEELVVEMLKPELKAWLDKNLPVLVKELVEIEIKKLVQNSKR-----    (122)
Rw    NFTDTIKDKKLDNNISSKNALEELVVEMLKPELKAWLDKNLPVLVKELVEIEIKKLVQNSKR-----    (132)
Hj    NFTDTIKDKKLDNNISSKNALEELVIEMLKPELKAWLDKNLSVLVKELVEIEIKKLVQNSKR-----    (132)
Ja    NFTDTIKDKKLDNNISSKNALEELVIEMLKPELKAWLDKNLSVLVKELVEIEIKKLVQNSKR-----    (132)
Co    NFTDTIKDKKLDNNISSKNALEELVIEMLKPELKAWLDKNLPVLVKELVEIEIKKLVQNSKR-----    (132)
Si    NFTDTIKDKKLDNNISSKNALEELVIEMLKPELKAWLDKNLPVLVKELVEIEIKKLVQNSKR-----    (122)
Af    NFTDTIKDKKLDNNISSKNALEELVIEMLKPELKAWLDKNLPVLVKELVEIEIKKLVQNSKR-----    (132)
```

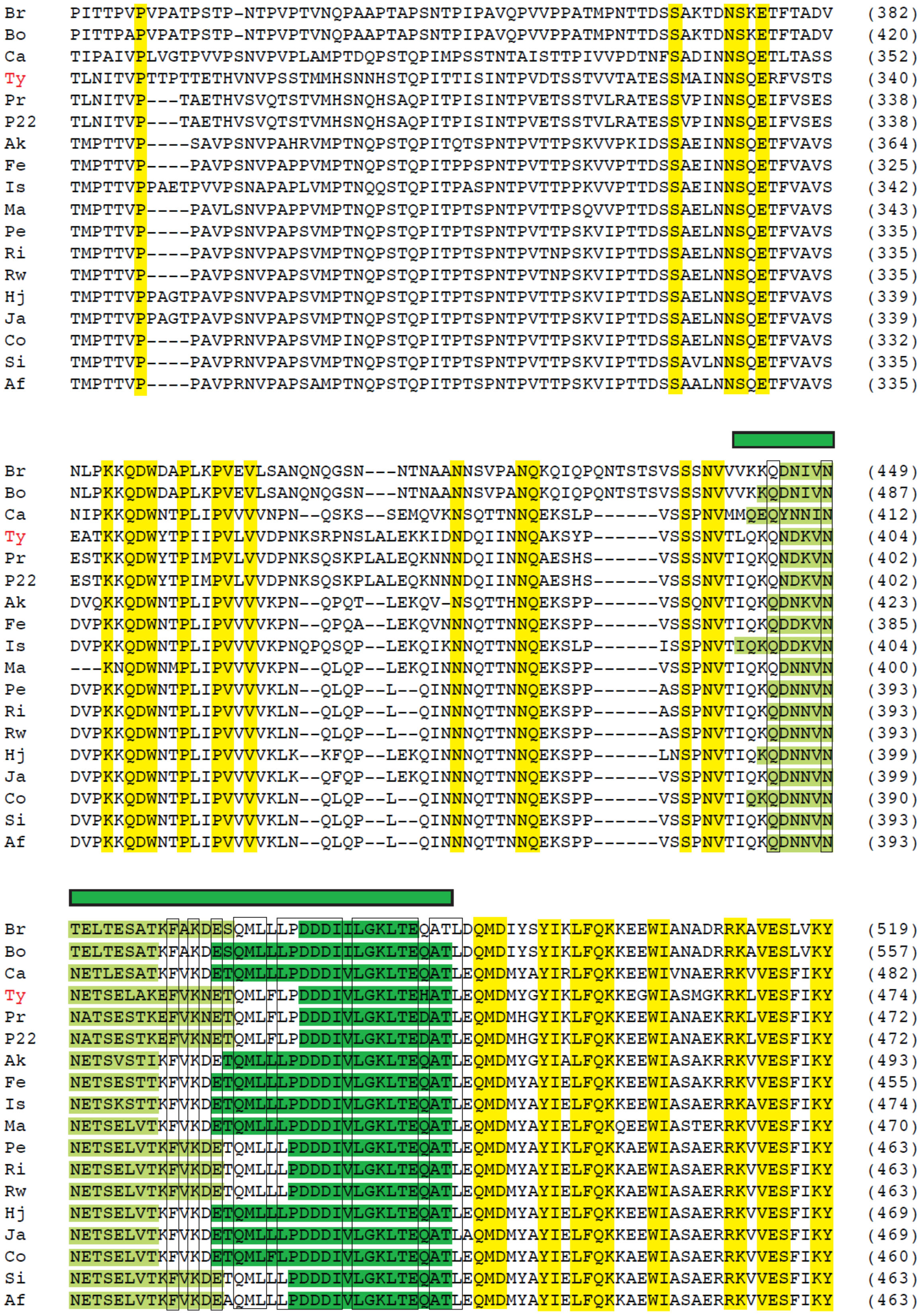


```
Br   --------MLNKLCRILFFISLSLATLQSYAAAPPPLPMPSQNSDIKS---KDEQPTSDSGSSMSIFDKI   (59)
Bo   --------MLNKLCRILFFISLSLATLQSYAAAPPPLPMPSQNSDIKS---KDEQPTSDSGSSMSIFDKI   (59)
Ca   --------MLNKLCKILFFIKLLLMTGQSY-AVPPTLPPSLPDVEVAT--TEDKEFRYNSDTSI--FDKF   (57)
Ty   --------MLNKLCEILFLINLLLVTGQGY-ASPPPLPQSLPIVAIDTPDKEDKDVSTNFNISF--FEKL   (59)
Pr   --------MLNKLCDILFLINLLLVTVQGY-ASPPPLPPSLPIIAVDT---TDKNISTNSNISF--FEKF   (56)
P22  --------MLNKLCDILFLINLLLVTVQGY-ASPPPLPPSLPIIAVDT---TDKNISTNSNISF--FEKF   (56)
Ak   --------MLNKSCKILFFINLLLVVVQSY-ASPPPLPTSLPAAEVDT---KDKDVSSNSDISF--FDKF   (56)
Fe   --------MLNKLCKILFFINLLLVAVQSY-ASPPP---SLPAAEIDT---KDKDVSSNSDISF--FDKF   (53)
Is   --------MLNKLCKILFFINLLLVAVQSY-ASPPSLPPSLPVAEVDT---KDKDVSSNSDISF--FDKI   (56)
Ma   MNYLIRNYMLNKLCKILFFINLLLVTVQSY-ASPPPLPPSLPVVAVDT---KDKDVRSSADISF--FDKF   (64)
Pe   --------MLNKLCKILFFINLLLVIVQSY-ASPPPLPPSLPAAEVDT---KDKDVRSSADISF--FDKF   (56)
Ri   --------MLNKLCKILFFINLLLVIVQSY-ASPPPLPPSLPAAEVDT---KDKDVRASADISF--FDKF   (56)
Rw   --------MLNKLCKILFFINLLLVIVQSY-ASPPPLPPSLPAAEVDT---KDKDVRASADISF--FDKF   (56)
Hj   --------MLNKLCKILFFINLLLVIVQSY-ASPPPLPPSLPAVEVDT---KYKDVRSSADVSF--FDKF   (56)
Ja   --------MLNKLCKILFFINLLLVIVQSY-ASPPPLPPSLPTAEVDT---KDKDVRSSADVSF--FDKF   (56)
Co   --------MLNNLCKILFFINLLLVIVQSY-ASPPPLPPSLPAAAVDT---KDKDVRSSADISF--FDKF   (56)
Si   --------MLNKLCKILFFINLLLVIVQSY-ASPPPLPPSLPAAEVDT---KDKDVRSSANISF--FDKF   (56)
Af   --------MLNKLCKILFFINLLLVIVQSY-ASPPPLPPSLPAAEVDT---KDKDVRSSADISF--FDKF   (56)

Br   KQFFHKSPKKKPLPKPQAQPDKPSDKLVSQEPNKNEVQLPSANNEVHQTNMNLASHND------------  (117)
Bo   KQFFHKSPKKKPLPKPQAQPDKPSDKLVSQEPNKNEVQLPSANNEVHQTNMNLASHNDTNSEKEASEPFI  (129)
Ca   KQFFSK-PQKKYISPKSTNEQ---------------------TKTADQEEPKL-SQEP------------   (92)
Ty   KQFFSK-QKKTNISSQQEKGQ---------------------TKAIHKESQQIDPREL------------   (95)
Pr   KQFFSK-QKKKNISSQHEQEQ---------------------TKAIHQESQQIDSREL------------   (92)
P22  KQFFSK-QKKKNISSQHEQEQ---------------------TKAIHQESQQIDSREL------------   (92)
Ak   KQFFSK-SKKKNIHTKQPNEQ---------------------TKAAPQEEPKLASHEH------------   (92)
Fe   KQFFSK-SKKKNISPKQPNEQ---------------------TKAANQEEPKLASQEH------------   (89)
Is   KQFFSK-SKKKNIPPKQANEQ---------------------TKAAHKEEPKLASQES------------   (92)
Ma   KQFFSK-SKKKDIPPKQANEQ---------------------TKAAHQEELKLASQEP------------  (100)
Pe   KQFFSK-PKKKDIPPKQANEQ---------------------TKAAHQEEPKLASQEP------------   (92)
Ri   KQFFSK-PKKKDIPPKQANEQ---------------------TKAAHQEEPKLASQEP------------   (92)
Rw   KQFFSK-PKKKDIPPKQANEQ---------------------TKAAHQEEPKLASQEP------------   (92)
Hj   KQFFSK-PKKKDIPPKQANEQ---------------------TKAAHQEEPKLASQEP------------   (92)
Ja   KQFFSK-PKKKDIPPKQANEQ---------------------TKAAHQEEPKLASQEP------------   (92)
Co   KQFFSK-PKKKDIPPKQASEQ---------------------TKAAHQEEPKLASQEP------------   (92)
Si   KQFFSK-PKKKDIPPKQASEQ---------------------TKAAHQEEPKLASKGP------------   (92)
Af   KQFFSK-PKKKDIPPKHASEQ---------------------TKAAHQEEPKLASQEP------------   (92)

Br   --------------------------TNSEKEASEPFIDMGSATLPSASNQ-------------------  (142)
Bo   DMGSATLPSASNQSHQANANLASHSGTNSEKEASEPFIDMGSATLPSASNQ-------------------  (180)
Ca   ----------------------------NENEHADSFMDISNVVLPSAASNNDVYTNT------------  (122)
Ty   ----------------------------NENEQSEPFIDVGSVILPSVAN---IHSKF------------  (122)
Pr   ----------------------------NENEQSEPFIDFGSTILPSVASNY-IDSKA------------  (121)
P22  ----------------------------NENEQSEPFIDFGSTILPSVASNY-IDSKA------------  (121)
Ak   ----------------------------TEDDVSKPFIDVGNTALPSAIASY-EHENGVNLAFHDIQESS  (133)
Fe   ----------------------------TEADASEPFIDTGNTALPSVTASN------------------  (113)
Is   ----------------------------NENEQAEPFIDVGNTALPSATAIN------------------  (116)
Ma   ----------------------------NDNEQAEPFIDVGNTALPSATSDN-VHTNP------------  (129)
Pe   ----------------------------NDNEQAEPFIDVGNTALPSATSDN-VHTNP------------  (121)
Ri   ----------------------------NDNEQAEPFIDVGNTALPSATSDN-VHTNP------------  (121)
Rw   ----------------------------NDNEQAEPFIDVGNTALPSATSDN-VHTNP------------  (121)
Hj   ----------------------------NDNEQAEQFIDVGNTALPSATSDN-VHANS------------  (121)
Ja   ----------------------------NDNEQAEPFIDVGNTALPSATSDN-VHANS------------  (121)
Co   ----------------------------NDNEQAEPFIDVGNTALPSATSDN-VHTNP------------  (121)
Si   ----------------------------NDNEQAEPFIDVGNTALPSATSDN-VHTNP------------  (121)
Af   ----------------------------NDNEQAEPFIDVGNTVLPSATSDN-VHTNP------------  (121)
```

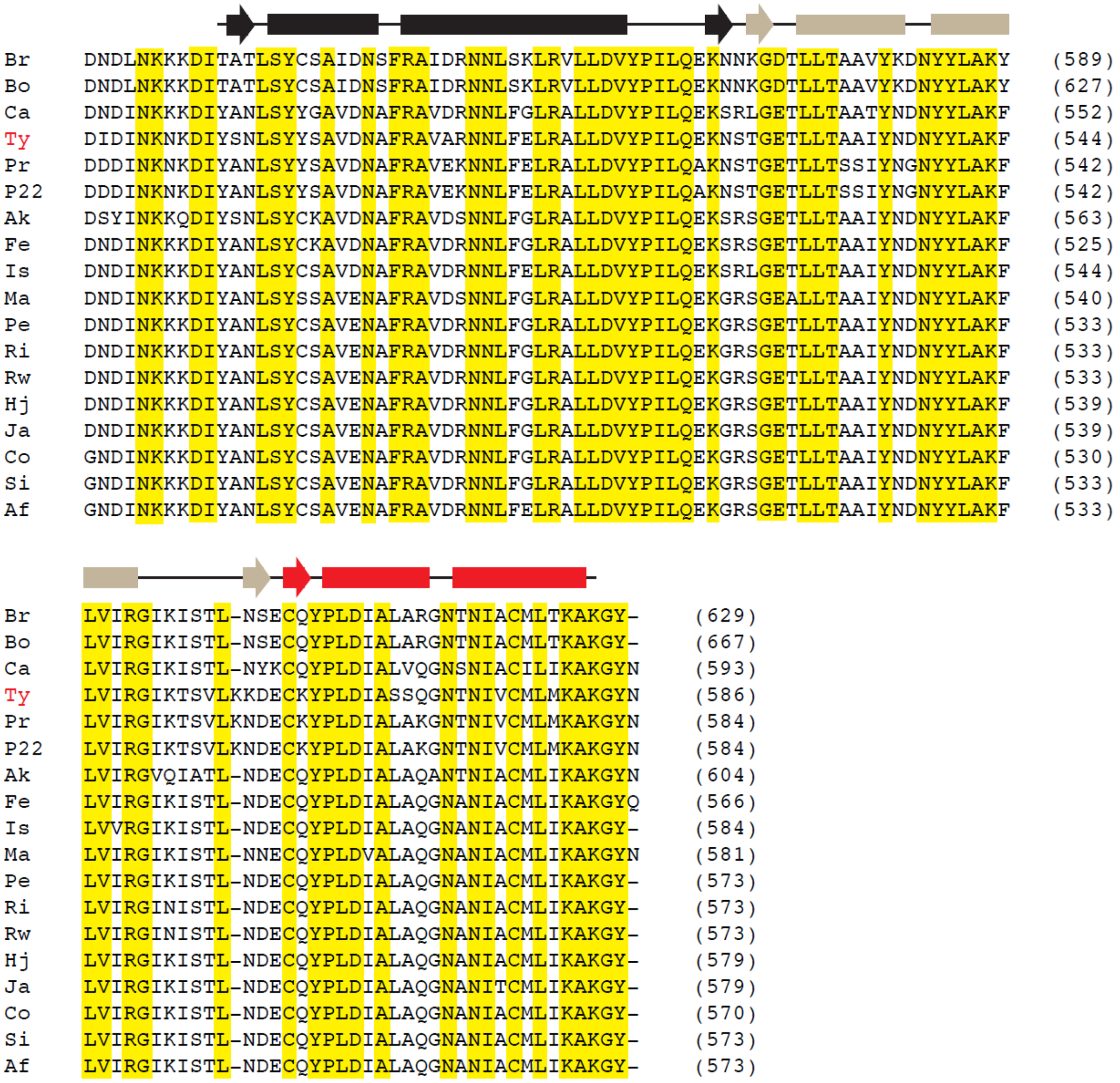
```
Br     -----------------------THQTNTNLASHS---------DTNSEKEASEPFIDMGNATLPSA--S   (178)
Bo     -----------------------THQTNTNLASHS---------DTNSEKEASEPFIDMGNATLPSA--S   (216)
Ca     -----------------------NHESSVNLASYDNIPDIQVMQQEPSESEIAETFIDIGSTKLPSA-TS   (168)
Ty     -----------------------EYENSTNLASNYNTQDMQVTQQELNSTEASESFIDIGNTKFISA-TN   (168)
Pr     -----------------------EYENSTNLAASYNTQDIQVKQQEFDPSEASEP-IDIGNTKFTSA-TN   (166)
P22    -----------------------EYENSTNLAASYNTQDIQVKQQEFDPSEASEP-IDIGNTKFTSA-TN   (166)
Ak     ETEASKPFIDVGNTALPSAIASYEHENGVNLAFHD--------IQESSETEASKPFIDVGNTALPSAIAS   (195)
Fe     -----------------------EHENSVNLASHD--------IQESNETEASKPFIDVGNTALPSA-AS   (151)
Is     -----------------------EHESSVNLASHDNTPDKQVTQQETNESETSEPFIDIGSAKLPSA-TS   (162)
Ma     -----------------------SHESSVNLASHD--------MQESNEAEASEPFIDIGSATLPSV-TS   (167)
Pe     -----------------------SHESSVNLASHD--------IQESNEAEASEPFIDIGSATLPSV-TS   (159)
Ri     -----------------------SHESSVNLASHD--------IQESNEAEASEPFIDIGSATLPSV-TS   (159)
Rw     -----------------------SHESSVNLASHD--------IQESNEAEASEPFIDIGSATLPSV-TS   (159)
Hj     -----------------------SHESSVNLAYHA--------IQESNEAEASEPFIDIGSATLPSV-TS   (159)
Ja     -----------------------SHESSVNLAYHD--------IQESNEAEASEPFIDIGSATLPSV-TS   (159)
Co     -----------------------SHESSVNLASHD--------IQESNEAEASEPFIDIGSATLPSV-TS   (159)
Si     -----------------------SHESSVNLASHE--------IQESNEVEVSESFIDIGSATLPSV-TS   (159)
Af     -----------------------SHESSVNLASHD--------IQESNEAEVSEPFIDIGSATLPSV-TS   (159)
```

```
Br     NQTHQANANLASHDDTKDGAASSEKEANSTPLPNTANNEEAQPELKVAGSLISNPPLRPGSYVVPPAPRT   (248)
Bo     NQTHQANANLASHDDTKDGAASSEKEANSTPLPNTANNEEAQPELKVAGSLISNPPLRPGSYVVPPAPRT   (286)
Ca     DEMRDAEIAANVH-DEN--LASNIITRNNIPRP-----------------MVSIPLARPGSYVVPPSPPV   (218)
Ty     NEMYKEAVLPNDYKKTN--LDFNIITL-NVLKP-----------------V--IPTSQDMMDVVPLQQPV   (216)
Pr     HEMYKEAVSSND-KETN--LTSNIITP-NVPSP-----------------VISIPTAQDVNYVVPSQQSV   (215)
P22    HEMYKEAVSSND-KETN--LTSNIITP-NVPSP-----------------VISIPTAQDVNYVVPSQQSV   (215)
Ak     YEVHEEEVASNEHKDTN--LSSNIITP-NVPRP-----------------IVSMPPAQAGSYVVPHRPV   (245)
Fe     NDVH---TNTEHENSTN--LASNIITP-NVPRP-----------------IVSMPPAQAGSYVVPPQRPV   (198)
Is     NEMHEEEVDSNEHKDTN--LASNIITP-NVPRP-----------------IVSMPPAQAGSYIVPPQRPV   (212)
Ma     NEMHEEQVASNEHSDTN--LASNIITP-NVPRP-----------------IVSMPPAQAGSYVVPPQRPV   (217)
Pe     NEMHEAQVASNEHNDTN--LASNIIIP-NVPRP-----------------IVSMPPAQAGSYVVPPQRPV   (209)
Ri     NEMHEAQVASNEHNDTN--LASNIIIP-NVPRP-----------------IVSMPPAQAGSYLVPPQRPV   (209)
Rw     NEMHEAQVASNEHNDTN--LASNIIIP-NVPRP-----------------IVSMPPAQAGSYLVPPQRPV   (209)
Hj     NEMHEAQVASNEHNDTN--LASNIITP-NVSRP-----------------IVSMPPAQAGSYVVPPQRPV   (209)
Ja     NEMHEAQVASNEHNDTN--LVSNIITP-NVPRP-----------------IVSMPPAQAGSYVVPPQRPV   (209)
Co     NEMHEAQVASNEHNDTN--LASNIIIP-NVPRP-----------------IVSMPPAQAESYVVPPQRPV   (209)
Si     NEMHEAQVASNEHNDTN--LASNIIIP-NVPRP-----------------IVSMPPAQAESYVVPPQRPV   (209)
Af     NEMHEAQVASNEHNDTN--LASNIIIP-NVPRP-----------------IVSMPPAQAESYVVPPQRPV   (209)
```

```
Br     QVYQPIALPPTHQY-IKLTPPPEANEQQDNVTAPPPP---QVVAPAPTVMPATPIPVVNQPTASD-VVTP   (313)
Bo     QVYQPIALPPTHQY-IKLTPPPEANEQQDNVTAPPPP---QVVAPAPTVMPATPIPVVNQPTASD-VVTP   (351)
Ca     QIYKPTNLLAVHKKHILLNPPPDVHKTQESIIPIAPPATSSI-----PNMPAISLPAVSTPVTQD-TNPS   (282)
Ty     QIYKPTNLTSI--P-KLFNHDTNLNNVEKNLES-------TM-----SNMTTISPNMLSVPTTQD-TIPT   (270)
Pr     QIYKPTNLTSIRNP-IPLNHHTDLNKVEKNLES-------TI-----SNMTTIPTNMVSVPSIQD-TIQT   (271)
P22    QIYKPTNLTSIRNP-IPLNHHTDLNKVEKNLES-------TI-----SNMTTIPTNMVSVPSIQD-TIQT   (271)
Ak     QIYKPTNLLSVHRP-IPLNTPPDAYKEAESVAPQ------SI-----PNMPAVSPPVI-----QDITTPS   (298)
Fe     QIYKPTNLPPVHKP-IPLNPSPEANKEEESVAPIAPQ---SI-----PNMPAVSPPVVSPPVIQDTTTPS   (259)
Is     QIYKPTNLPPVHKP-IPLNPPPDANKEEESVAPIAPP---SI-----PNMPAVSPPVVSPPVTQD-ITPS   (272)
Ma     QIYKPTNLPPVHKP-IPLNPPPDANKEEESVAPIAPP---SI-----PNMPAVSPPVVSPPVTQD-TTSS   (277)
Pe     QIYKPTNLPPVHKP-IPLNPPPDANKEEESVAPIAPP---SI-----TNIPAVSPPVVSPPVTQD-NTSS   (269)
Ri     QIYTPTNLPPVHKP-IPLNPPPDANKEEESVAPIAPP---SI-----TNIPAVSPPVVSPPVTQD-NTSS   (269)
Rw     QIYTPTNLPPVHKP-IPLNPPPDANKEEESVAPIAPP---SI-----TNIPAVSPPVVSPPVTQD-NTSS   (269)
Hj     QIYKPTNLPPVHKP-IPLNPPPDANKEEESVAPIAPP---SI-----PNMPAVSPPVVNPPVTQD-TTSS   (269)
Ja     QIYKPTNLPPVHKP-IPLNPTPDANKEEESVAPIAPP---SI-----PNMPAVSPPVVNPPVTQD-TTSS   (269)
Co     QIYKPTNLPPVHKP-IPLNPPPDANKEEESA---ALP---SI-----TNIPAVSPPVVSPPVTQD-NTSS   (266)
Si     QIYKPTNLPPVHKP-IPFNPPPDANKEEESVAPIAPP---SI-----TNIPAVSPPVVSPPVTQD-NTSS   (269)
Af     QIYKPTNLPPVHKP-IPLNPPPDANKEEESVAPIAPP---SI-----TNIPAVSLPVVSPPVTQD-NTSS   (269)
```



```
Br     PITTPVPVPATPSTP-NTPVPTVNQPAAPTAPSNTPIPAVQPVVPPATMPNTTDSSAKTDNSKETFTADV   (382)
Bo     PITTPAPVPATPSTP-NTPVPTVNQPAAPTAPSNTPIPAVQPVVPPATMPNTTDSSAKTDNSKETFTADV   (420)
Ca     TIPAIVPLVGTPVVPSNVPVPLAMPTDQPSTQPIMPSSTNTAISTTPIVVPDTNFSADINNSQETLTASS   (352)
Ty     TLNITVPTTPTTETHVNVPSSTMMHSNNHSTQPITTISINTPVDTSSTVVTATESSMAINNSQERFVSTS   (340)
Pr     TLNITVP---TAETHVSVQTSTVMHSNQHSAQPITPISINTPVETSSTVLRATESSVPINNSQEIFVSES   (338)
P22    TLNITVP---TAETHVSVQTSTVMHSNQHSAQPITPISINTPVETSSTVLRATESSVPINNSQEIFVSES   (338)
Ak     TMPTTVP----SAVPSNVPAHRVMPTNQPSTQPITQTSPNTPVTTPSKVVPKIDSSAEINNSQETFVAVS   (364)
Fe     TMPTTVP----PAVPSNVPAPPVMPTNQPSTQPITPPSPNTPVTTPSKVVPTTDSSAEINNSQETFVAVS   (325)
Is     TMPTTVPPAETPVVPSNAPAPLVMPTNQQSTQPITPASPNTPVTTPPKVVPTTDSSAEINNSQETFVAVS   (342)
Ma     TMPTTVP----PAVLSNVPAPPVMPTNQPSTQPITPTSPNTPVTTPSQVVPTTDSSAELNNSQETFVAVS   (343)
Pe     TMPTTVP----PAVPSNVPAPSVMPTNQPSTQPITPTSPNTPVTTPSKVIPTTDSSAELNNSQETFVAVS   (335)
Ri     TMPTTVP----PAVPSNVPAPSVMPTNQPSTQPITPTSPNTPVTNPSKVIPTTDSSAELNNSQETFVAVS   (335)
Rw     TMPTTVP----PAVPSNVPAPSVMPTNQPSTQPITPTSPNTPVTNPSKVIPTTDSSAELNNSQETFVAVS   (335)
Hj     TMPTTVPPAGTPAVPSNVPAPSVMPTNQPSTQPITPTSPNTPVTTPSKVIPTTDSSAELNNSQETFVAVS   (339)
Ja     TMPTTVPPAGTPAVPSNVPAPSVMPTNQPSTQPITPTSPNTPVTTPSKVIPTTDSSAELNNSQETFVAVS   (339)
Co     TMPTTVP----PAVPRNVPAPSVMPINQPSTQPITPTSPNTPVTTPSKVIPTTDSSAELNNSQETFVAVS   (332)
Si     TMPTTVP----PAVPRNVPAPSVMPTNQPSTQPITPTSPNTPVTTPSKVIPTTDSSAVLNNSQETFVAVS   (335)
Af     TMPTTVP----PAVPRNVPAPSAMPTNQPSTQPITPTSPNTPVTTPSKVIPTTDSSAALNNSQETFVAVS   (335)

Br     NLPKKQDWDAPLKPVEVLSANQNQGSN---NTNAANNSVPANQKQIQPQNTSTSVSSSNVVVKKQDNIVN   (449)
Bo     NLPKKQDWDAPLKPVEVLSANQNQGSN---NTNAANNSVPANQKQIQPQNTSTSVSSSNVVVKKQDNIVN   (487)
Ca     NIPKKQDWNTPLIPVVVVNPN--QSKS--SEMQVKNSQTTNNQEKSLP------VSSPNVMMQEQYNNIN   (412)
Ty     EATKKQDWYTPIIPVLVVDPNKSRPNSLALEKKIDNDQIINNQAKSYP------VSSSNVTLQKQNDKVN   (404)
Pr     ESTKKQDWYTPIMPVLVVDPNKSQSKPLALEQKNNNDQIINNQAESHS------VSSSNVTIQKQNDKVN   (402)
P22    ESTKKQDWYTPIMPVLVVDPNKSQSKPLALEQKNNNDQIINNQAESHS------VSSSNVTIQKQNDKVN   (402)
Ak     DVQKKQDWNTPLIPVVVVKPN--QPQT--LEKQV-NSQTTHNQEKSPP------VSSQNVTIQKQDNKVN   (423)
Fe     DVPKKQDWNTPLIPVVVVKPN--QPQA--LEKQVNNNQTTNNQEKSPP------VSSSNVTIQKQDDKVN   (385)
Is     DVPKKQDWNTPLIPVVVVKPNQPQSQP--LEKQIKNNQTTNNQEKSLP------ISSPNVTIQKQDDKVN   (404)
Ma     ---KNQDWNMPLIPVVVVKPN--QLQP--LEKQINNNQTTNNQEKSPP------VSSPNVTIQKQDNNVN   (400)
Pe     DVPKKQDWNTPLIPVVVVKLN--QLQP--L--QINNNQTTNNQEKSPP------ASSPNVTIQKQDNNVN   (393)
Ri     DVPKKQDWNTPLIPVVVVKLN--QLQP--L--QINNNQTTNNQEKSPP------ASSPNVTIQKQDNNVN   (393)
Rw     DVPKKQDWNTPLIPVVVVKLN--QLQP--L--QINNNQTTNNQEKSPP------ASSPNVTIQKQDNNVN   (393)
Hj     DVPKKQDWNTPLIPVVVVKLK--KFQP--LEKQINNNQTTNNQEKSPP------LNSPNVTIQKQDNNVN   (399)
Ja     DVPKKQDWNTPLIPVVVVKLK--QFQP--LEKQINNNQTTNNQEKSPP------VSSPNVTIQKQDNNVN   (399)
Co     DVPKKQDWNTPLIPVVVVKLN--QLQP--L--QINNNQTTNNQEKSPP------VSSPNVTIQKQDNNVN   (390)
Si     DVPKKQDWNTPLIPVVVVKLN--QLQP--L--QINNNQTTNNQEKSPP------VSSPNVTIQKQDNNVN   (393)
Af     DVPKKQDWNTPLIPVVVVKLN--QLQP--L--QINNNQTTNNQEKSPP------VSSPNVTIQKQDNNVN   (393)

Br     TELTESATKFAKDESQMLLLPDDDIILGKLTEQATLDQMDIYSYIKLFQKKEEWIANADRRKAVESLVKY   (519)
Bo     TELTESATKFAKDESQMLLLPDDDIILGKLTEQATLDQMDIYSYIKLFQKKEEWIANADRRKAVESLVKY   (557)
Ca     NETLESATKFVKDETQMLLLPDDDIVLGKLTEQATLEQMDMYAYIRLFQKKEEWIVNAERRKVVESFIKY   (482)
Ty     NETSELAKEFVKNETQMLFLPDDDIVLGKLTEHATLEQMDMYGYIKLFQKKEGWIASMGKRKLVESFIKY   (474)
Pr     NATSESTKEFVKNETQMLFLPDDDIVLGKLTEDATLEQMDMHGYIKLFQKKEEWIANAEKRKLVESFIKY   (472)
P22    NATSESTKEFVKNETQMLFLPDDDIVLGKLTEDATLEQMDMHGYIKLFQKKEEWIANAEKRKLVESFIKY   (472)
Ak     NETSVSTIKFVKDETQMLLLPDDDIVLGKLTEQATLEQMDMYGYIALFQKKEEWIASAKRRKVVESFIKY   (493)
Fe     NETSESTTKFVKDETQMLLLPDDDIVLGKLTEQATLEQMDMYAYIELFQKKEEWIASAKRRKVVESFIKY   (455)
Is     NETSKSTTKFVKDETQMLLLPDDDIVLGKLTEQATLEQMDMYAYIELFQKKEEWIASAERRKAVESFIKY   (474)
Ma     NETSELVTKFVKDETQMLLLPDDDIVLGKLTEQATLEQMDMYAYIELFQKQEEWIASTERRKVVESFIKY   (470)
Pe     NETSELVTKFVKDETQMLLLPDDDIVLGKLTEQATLEQMDMYAYIKLFQKKAEWIASAERRKVVESFIKY   (463)
Ri     NETSELVTKFVKDETQMLLLPDDDIVLGKLTEQATLEQMDMYAYIELFQKKAEWIASAERRKVVESFIKY   (463)
Rw     NETSELVTKFVKDETQMLLLPDDDIVLGKLTEQATLEQMDMYAYIELFQKKAEWIASAERRKVVESFIKY   (463)
Hj     NETSELVTKFVKDETQMLLLPDDDIVLGKLTEQATLEQMDMYAYIELFQKKEEWIASAERRKVVESFIKY   (469)
Ja     NETSELVTKFVKDETQMLLLPDDDIVLGKLTEQATLAQMDMYAYIELFQKKEEWIASAERRKVVESFIKY   (469)
Co     NETSELVTKFVKDETQMLFLPDDDIVLGKLTEQATLEQMDMYAYIELFQKKAEWIASAERRKVVESFIKY   (460)
Si     NETSELVTKFVKDETQMLLLPDDDIVLGKLTEQATLEQMDMYAYIELFQKKAEWIASAERRKVVESFIKY   (463)
Af     NETSELVTKFVKDEAQMLLLPDDDIVLGKLTEQATLEQMDMYAYIELFQKKAEWIASAERRKVVESFIKY   (463)
```



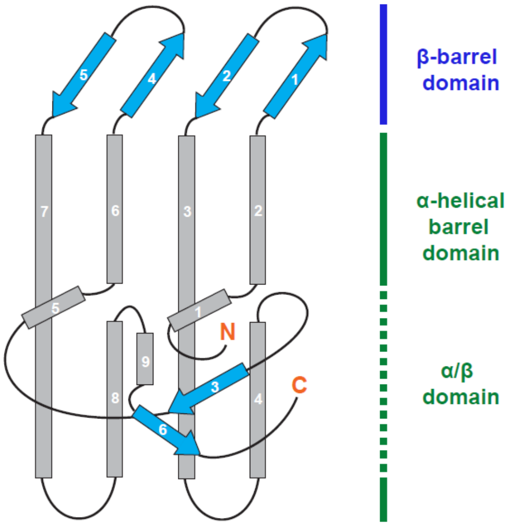
```
Br     DNDLNKKKDITATLSYCSAIDNSFRAIDRNNLSKLRVLLDVYPILQEKNNKGDTLLTAAVYKDNYYLAKY   (589)
Bo     DNDLNKKKDITATLSYCSAIDNSFRAIDRNNLSKLRVLLDVYPILQEKNNKGDTLLTAAVYKDNYYLAKY   (627)
Ca     DNDINKKKDIYANLSYYGAVDNAFRAVDRNNLFGLRALLDVYPILQEKSRLGETLLTAATYNDNYYLAKF   (552)
Ty     DIDINKNKDIYSNLSYYSAVDNAFRAVARNNLFELRALLDVYPILQEKNSTGETLLTAAIYNDNYYLAKF   (544)
Pr     DDDINKNKDIYANLSYYSAVDNAFRAVEKNNLFELRALLDVYPILQAKNSTGETLLTSSIYNGNYYLAKF   (542)
P22    DDDINKNKDIYANLSYYSAVDNAFRAVEKNNLFELRALLDVYPILQAKNSTGETLLTSSIYNGNYYLAKF   (542)
Ak     DSYINKKQDIYSNLSYCKAVDNAFRAVDSNNLFGLRALLDVYPILQEKSRSGETLLTAAIYNDNYYLAKF   (563)
Fe     DNDINKKKDIYANLSYCKAVDNAFRAVDRNNLFGLRALLDVYPILQEKSRSGETLLTAAIYNDNYYLAKF   (525)
Is     DNDINKKKDIYANLSYCSAVDNAFRAVDRNNLFELRALLDVYPILQEKSRLGETLLTAAIYNDNYYLAKF   (544)
Ma     DNDINKKKDIYANLSYSSAVENAFRAVDSNNLFGLRALLDVYPILQEKGRSGEALLTAAIYNDNYYLAKF   (540)
Pe     DNDINKKKDIYANLSYCSAVENAFRAVDRNNLFGLRALLDVYPILQEKGRSGETLLTAAIYNDNYYLAKF   (533)
Ri     DNDINKKKDIYANLSYCSAVENAFRAVDRNNLFGLRALLDVYPILQEKGRSGETLLTAAIYNDNYYLAKF   (533)
Rw     DNDINKKKDIYANLSYCSAVENAFRAVDRNNLFGLRALLDVYPILQEKGRSGETLLTAAIYNDNYYLAKF   (533)
Hj     DNDINKKKDIYANLSYCSAVENAFRAVDRNNLFGLRALLDVYPILQEKGRSGETLLTAAIYNDNYYLAKF   (539)
Ja     DNDINKKKDIYANLSYCSAVENAFRAVDRNNLFGLRALLDVYPILQEKGRSGETLLTAAIYNDNYYLAKF   (539)
Co     GNDINKKKDIYANLSYCSAVENAFRAVDRNNLFGLRALLDVYPILQEKGRSGETLLTAAIYNDNYYLAKF   (530)
Si     GNDINKKKDIYANLSYCSAVENAFRAVDRNNLFGLRALLDVYPILQEKGRSGETLLTAAIYNDNYYLAKF   (533)
Af     GNDINKKKDIYANLSYCSAVENAFRAVDRNNLFELRALLDVYPILQEKGRSGETLLTAAIYNDNYYLAKF   (533)

Br     LVIRGIKISTL-NSECQYPLDIALARGNTNIACMLTKAKGY-    (629)
Bo     LVIRGIKISTL-NSECQYPLDIALARGNTNIACMLTKAKGY-    (667)
Ca     LVIRGIKISTL-NYKCQYPLDIALVQGNSNIACILIKAKGYN    (593)
Ty     LVIRGIKTSVLKKDECKYPLDIASSQGNTNIVCMLMKAKGYN    (586)
Pr     LVIRGIKTSVLKNDECKYPLDIALAKGNTNIVCMLMKAKGYN    (584)
P22    LVIRGIKTSVLKNDECKYPLDIALAKGNTNIVCMLMKAKGYN    (584)
Ak     LVIRGVQIATL-NDECQYPLDIALAQANTNIACMLIKAKGYN    (604)
Fe     LVIRGIKISTL-NDECQYPLDIALAQGNANIACMLIKAKGYQ    (566)
Is     LVVRGIKISTL-NDECQYPLDIALAQGNANIACMLIKAKGY-    (584)
Ma     LVIRGIKISTL-NNECQYPLDVALAQGNANIACMLIKAKGYN    (581)
Pe     LVIRGIKISTL-NDECQYPLDIALAQGNANIACMLIKAKGY-    (573)
Ri     LVIRGINISTL-NDECQYPLDIALAQGNANIACMLIKAKGY-    (573)
Rw     LVIRGINISTL-NDECQYPLDIALAQGNANIACMLIKAKGY-    (573)
Hj     LVIRGIKISTL-NDECQYPLDIALAQGNANIACMLIKAKGY-    (579)
Ja     LVIRGIKISTL-NDECQYPLDIALAQGNANITCMLIKAKGY-    (579)
Co     LVIRGIKISTL-NDECQYPLDIALAQGNANIACMLIKAKGY-    (570)
Si     LVIRGIKISTL-NDECQYPLDIALAQGNANIACMLIKAKGY-    (573)
Af     LVIRGIKISTL-NDECQYPLDIALAQGNANIACMLIKAKGY-    (573)
```

Figure S2. Comparative sequence analysis of proteins encoded by (A) *RT0216*, (B) *RT0217* and (C) *RT0218* across 18 *Rickettsia* genomes. Sequences were obtained from the PATRIC website (4). The rickettsial species/strains are abbreviated as follows: Br, *R. bellii* str. RML369-C; Bo, *R. bellii* str. OSU 85 389; Ca, *R. canadensis* str. McKiel; Ty, *R. typhi* str. Wilmington; Pr, *R. prowazekii* str. Madrid E; P22, *R. prowazekii* str. Rp22; Ak, *R. akari* str. Hartford; Fe, *R. felis* str. URRWXCal2; IS, *Rickettsia* endosymbiont of *Ixodes scapularis*; Ma, *R. massilae* str. MTU5; Pe, *R. peacockii* str. Rustic; Ri, *R. rickettsii* str. Sheila Smith; Rw, *R. rickettsii* str. Iowa; Hj, *R.*

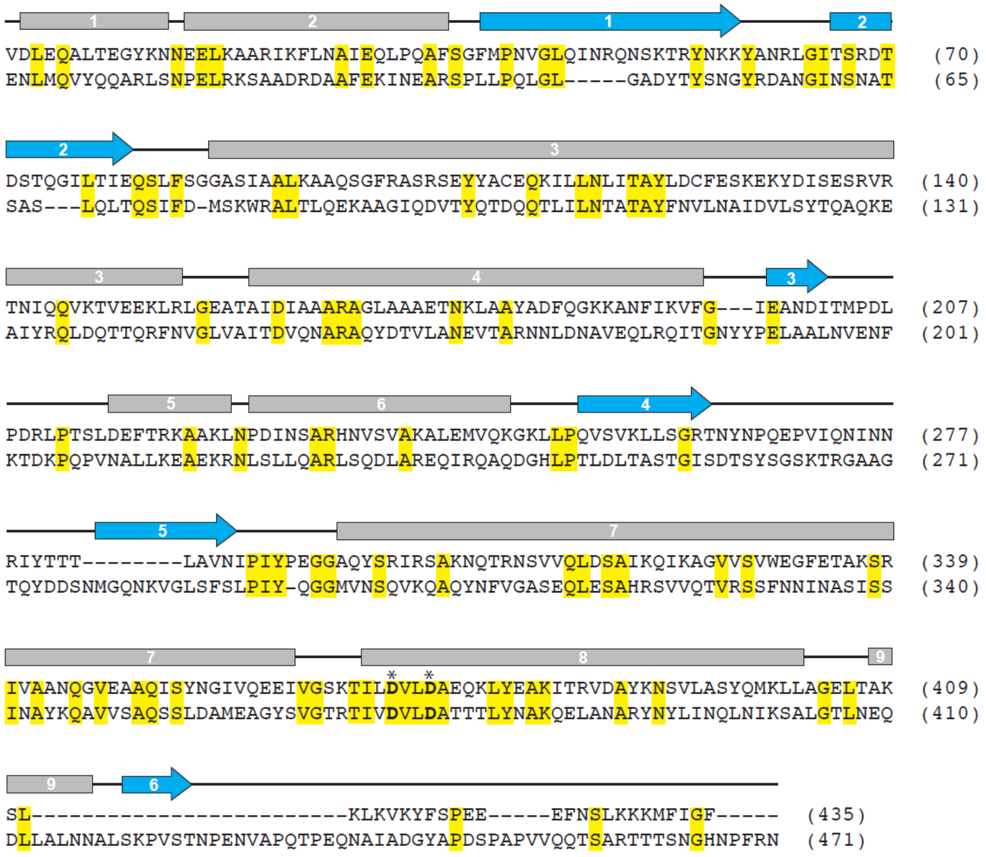
heilongjiangensis str. 054; Ja, *R. japonica* str. YH; Co, *R. conorii* str. Malish 7, Si, *R. sibirica* str. 246; Af, *R. africae* str. ESF-5. All alignments were constructed using MUSCLE v3.6 (3) with default parameters. Positions within the alignment that are highlighted yellow depict invariant residues across all sequences. Coordinates for each sequence (aa) are shown in parentheses at right. N-terminal signal sequences, as predicted with SignalP 4.0 (7), are shaded gray. (**A**) TolC protein alignment with secondary structure model shown at top (see Figure S4 for details). (**B**) Alignment of PopZ-like sequences with the conserved N- and C-terminal regions illustrated at top and colored blue and orange, respectively. The conserved regions are described in the text and shown in more detail in Figure S3. (**C**) RARP-1 protein alignment with repeat regions illustrated within the central domain. The repeat region proximal to the N-terminus is shown in blue and ranges from 2-4 repeats across rickettsial sequences (each repeat per sequence is shaded darker to demarcate the multiple repeats per sequence). The repeat region flanking the ANK domain is shown in green and contains two predicted repeats per sequence. Conserved residues within the repeat regions are boxed. Repeats were predicted using HHrepID v2.16.1 (1). The three ankyrin repeats within the ANK domain are illustrated as described in Figure 2 and Figure S1. NOTE: the region used to design the RARP-1 antibody (EKGQTKAIHKESQQIDPRE) is bolded (pos. 76-94).

Figure S3

A



B



```
Rt     VDLEQALTEGYKNNEELKAARIKFLNAIEQLPQAFSGFMPNVGLQINRQNSKTRYNKKYANRLGITSRDT    (70)
Ec     ENLMQVYQQARLSNPELRKSAADRDAAFEKINEARSPLLPQLGL-----GADYTYSNGYRDANGINSNAT    (65)

Rt     DSTQGILTIEQSLFSGGASIAALKAAQSGFRASRSEYYACEQKILLNLITAYLDCFESKEKYDISESRVR   (140)
Ec     SAS---LQLTQSIFD-MSKWRALTLQEKAAGIQDVTYQTDQQTLILNTATAYFNVLNAIDVLSYTQAQKE   (131)

Rt     TNIQQVKTVEEKLRLGEATAIDIAAARAGLAAAETNKLAAYADFQGKKANFIKVFG---IEANDITMPDL   (207)
Ec     AIYRQLDQTTQRFNVGLVAITDVQNARAQYDTVLANEVTARNNLDNAVEQLRQITGNYYPELAALNVENF   (201)

Rt     PDRLPTSLDEFTRKAAKLNPDINSARHNVSVAKALEMVQKGKLLPQVSVKLLSGRTNYNPQEPVIQNINN   (277)
Ec     KTDKPQPVNALLKEAEKRNLSLLQARLSQDLAREQIRQAQDGHLPTLDLTASTGISDTSYSGSKTRGAAG   (271)

Rt     RIYTTT--------LAVNIPIYPEGGAQYSRIRSAKNQTRNSVVQLDSAIKQIKAGVVSVWEGFETAKSR   (339)
Ec     TQYDDSNMGQNKVGLSFSLPIY-QGGMVNSQVKQAQYNFVGASEQLESAHRSVVQTVRSSFNNINASISS   (340)

                                      *  *
Rt     IVAANQGVEAAQISYNGIVQEEIVGSKTILDVLDAEQKLYEAKITRVDAYKNSVLASYQMKLLAGELTAK   (409)
Ec     INAYKQAVVSAQSSLDAMEAGYSVGTRTIVDVLDATTTLYNAKQELANARYNYLINQLNIKSALGTLNEQ   (410)

Rt     SL-------------------------KLKVKYFSPEE-----EFNSLKKKMFIGF-----   (435)
Ec     DLLALNNALSKPVSTNPENVAPQTPEQNAIADGYAPDSPAPVVQQTSARTTTSNGHNPFRN   (471)
```

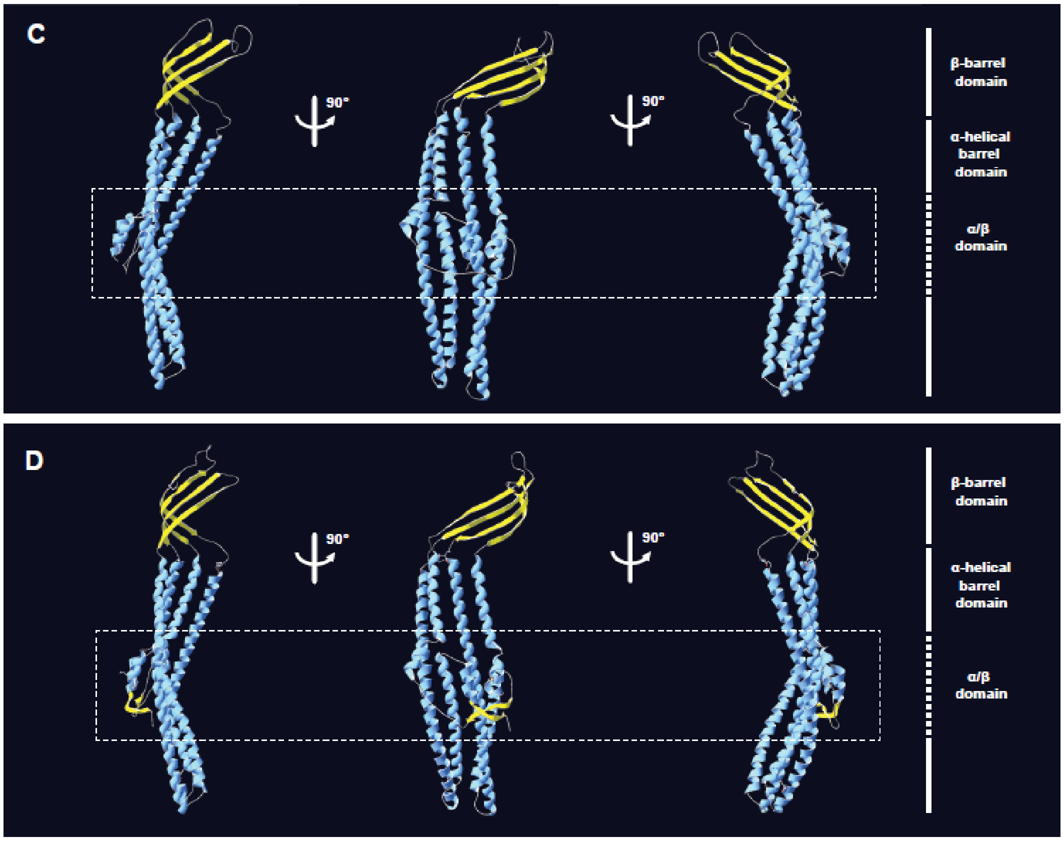


Figure S3. **Bioinformatics analysis of the *R. typhi* TolC.** (**A**) Diagram depicting the secondary structure of *Escherichia coli* TolC (5). The gray and blue bars correspond to the α-helices and β-strands, respectively. (**B**) Multiple sequence alignment of *R. typhi* (Rt, YP_067180) and *E. coli* (Ec, ZP_04872262) TolC proteins. Structural model at top depicts secondary structure shown in **A**. Conserved positions within the alignment are highlighted yellow. Coordinates for each sequence (aa) are shown in parentheses at right. (**C, D**) Structural modeling of *R. typhi* TolC (shown as a monomer, **C**) with the TolC monomer of *E. coli* (pdb 1ek9C, **D**). The α-helices are shown in blue, with β-strands colored yellow. The dashed box depicts the α/β domain that encompasses β-strands B3 and B6, which are absent in the *R. typhi* TolC structure prediction.

Figure S4

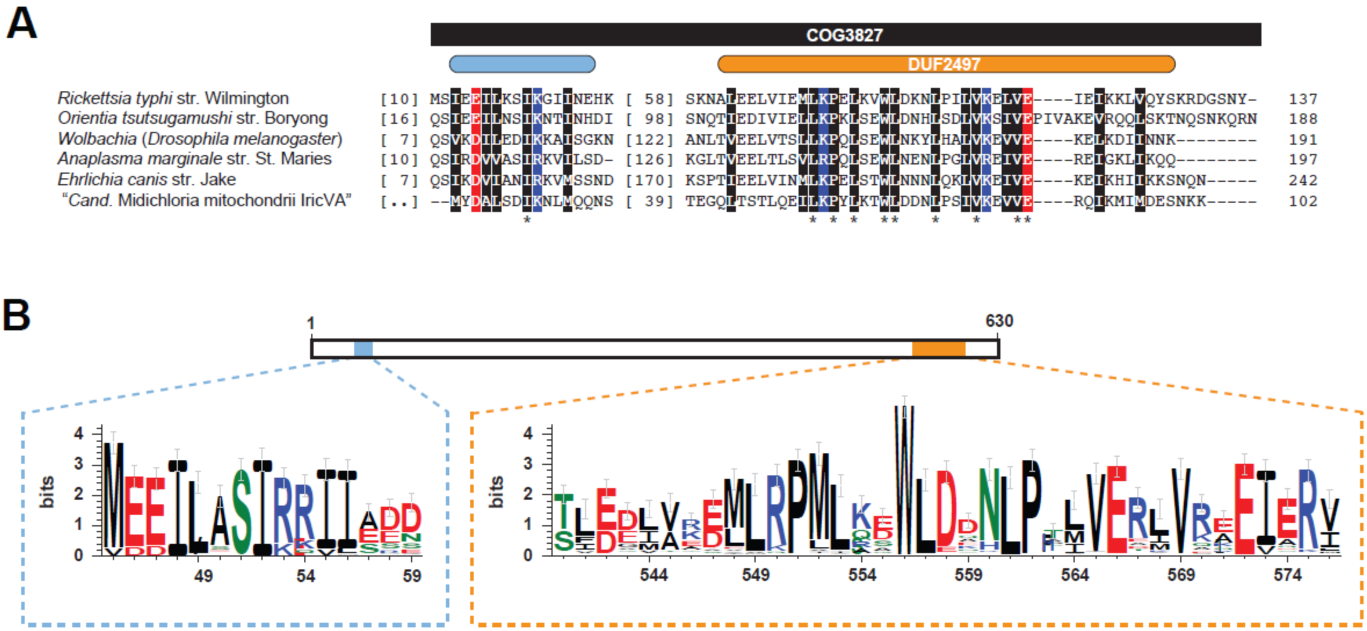


Figure S4. *R. typhi RT0217* encodes a putative PopZ homolog. (**A**) Multiple sequence alignment of six *Rickettsiales* PopZ-like proteins: HP RT0217 (*Rickettsia typhi* str. Wilmington, YP_067181); HP OTBS_0399 (*Orientia tsutsugamushi* str. Boryong, YP_001248323); HP WD0069 (*Wolbachia* endosymbiont of *Drosophila melanogaster*, NP_965897); HP AM1025 (*Anaplasma marginale* str. St. Maries, YP_154161); HP Ecaj_0824 (*Ehrlichia canis* str. Jake, YP_303453); HP midi_00757 ("*Candidatus* Midichloria mitochondrii str. IricVA", YP_004679733). See text for alignment details. NOTE: genes encoding PopZ-like sequences were not detected in *Neorickettsia* genomes. The conserved N-terminal (light-blue) and C-terminal (orange) domains are illustrated above the alignment. Invariant residues are depicted with an asterisk below the alignment. Amino acid conservation is shown for polar (green), non-polar (black), acidic (red), and basic (blue) residues. (**B**) A more extensive search for potential HP RT0217 homologs revealed 122 sequences within alphaproteobacterial genomes, with robust sequence comparison illustrating the conserved small N-terminal and C-terminal regions within an alignment encompassing extraordinary variability in sequence length and composition. Schema depicts multiple sequence alignment (1-630 positions) with conserved N- and C-terminal domains and sequence conservation across 122 alphaproteobacterial PopZ-like sequences shaded as in **A**. The conserved regions are illustrated with sequence logos (2), with amino acid classes colored as in **A**.

Figure S5

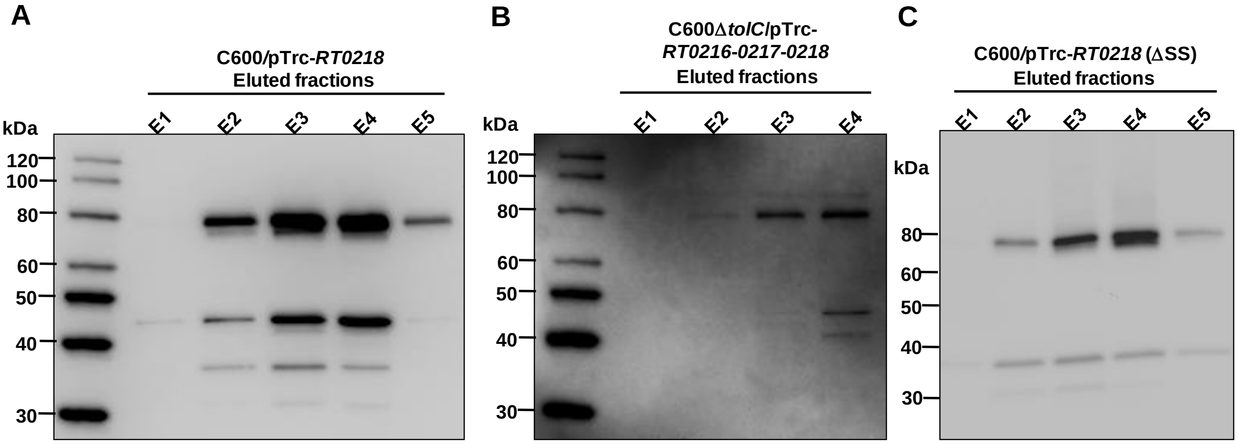


Figure S5. **Recombinant protein expression and purification of RARP-1**. *E. coli* C600/pTrc-*RT0218* (A), *E. coli* C600Δ*tolC*/pTrc-*RT0216-RT0217-RT0218* (B), and *E. coli* Top10 cells harboring pTrc-*RT0218* ΔSS (C) were induced by addition of IPTG to a final concentration of 1mM. The bacterial cells were harvested by centrifugation at 8,000 xg for 20 min and lysed with french press cell disrupter. The cleared cell lysate with recombinant full-length RARP-1 and RARP-1ΔSS was incubated with nickel chelated agarose and recombinant protein was purified under native conditions. The purified proteins were separated by SDS-PAGE following dialysis and concentrating. The recombinant full-length RARP-1 and RARP-1 ΔSS (~80kDa) was identified by immunoblotting using anti-RT0218 antibodies. The purity of the protein was assessed with coomassie stain (data not shown).

TABLES

Table S1. Primers used in the study.

Primer	Sequence (5’ ⟶ 3’)
RT0218-SS F	ATG TTA AAT AAA TTA TGT GAG ATA
RT0218-R2	ATT ATA CCC CTT AGC TTT CAT TAA
RT0218-F1	TCG CCT CCG CCA TTA CCA CAA TCA
RT0218-R3	AAT ATC TTT ATT TTT ATT GAT ATC
RT0216-F2	ATG AGT AAA TTA ACT ACA TTT ATT
RT0216-F1	TCC GCA AGA GCC AGT TAT TC
RT0216-R1	AAT CCT TCC CAC ACG CTA AC
RT0218 RT-F1	TCA TTC TAC ACA GCC TAT AAC GAC
RT0218 RT-R1	TGG GAT CAA CTA CAA GAA CTG G
RT0218 RT-R2	ACT AAC ATC CTT ATC TTC CTT ATC TGG
AZ5194-F	TCT ATT ATT GGT GAC AGG GC
AZ5195-R	CAT CCT TAT CTT CCT TAT CTG G
AZ4923	CTC CTG CCT TAG AAT CCA ACC C
AZ4924	TTC CTT TAC ACT ATG CTT TTC ACC AG
AZ2249-F	ACA ATC TGT GTG GGC ACT CG
AZ2248-R	GGC TGA AAA TCT TCT CTC ATC CGC

Table S2. RT0218 homologs of *Rickettsia* spp.

Accession number	Protein Homologue/Locus tag	Rickettsial species	Strain	Bit score	Identities
YP_067182.1	hypothetical protein RT0218	*Rickettsia typhi*	Wilmington	1192	
NP_220612.1	hypothetical protein RP226	*Rickettsia prowazekii*	Madrid E	827	79%
ZP_04699593.1	ankyrin repeat-containing domain protein	*Rickettsia endosymbiont of Ixodes scapularis*	-	647	63%
YP_004764027.1	Rh054_01760	*Rickettsia heilongjiangensis*	054	586	59%
YP_002844975.1	Ankyrin repeat protein	*Rickettsia africae*	ESF-5	585	58%
YP_247003.1	ankyrin repeat-containing protein RF_0987	*Rickettsia felis*	URRWXCal 2	580	59%
YP_002916264.1	hypothetical protein RPR_01695	*Rickettsia peacockii*	Rustic	578	59%
YP_001499117.1	ankyrin repeat-containing protein	*Rickettsia massiliae*	MTU5	578	60%
ZP_00142229.1	hypothetical protein	*Rickettsia sibirica*	246	576	58%
YP_001491992.1	threonyl-tRNA synthetase	*Rickettsia canadensis*	McKiel	577	54%
NP_359945.1	hypothetical protein RC0308	*Rickettsia conorii*	Malish 7	573	59%

YP_001494426.1	hypothetical protein A1G_01760	*Rickettsia rickettsii*	Sheila Smith	568	58%
YP_001649675.1	hypothetical protein RrIowa_0373	*Rickettsia rickettsii*	Iowa	568	58%
YP_001493155.1	ankyrin repeat-containing protein	*Rickettsia akari*	Hartford	558	55%
YP_537401.1	ankyrin repeat-containing protein	*Rickettsia bellii*	RML369-C	347	39%
YP_001496681.1	ankyrin repeat-containing protein	*Rickettsia bellii*	OSU 85-389	340	43%

Table S3. TolC homologs of *Rickettsia* spp.

Accession number	Protein Homologue	*Rickettsial species*	Strain	Bit score	Locus tag of homolog
YP_067180.1	Outer membrane protein TolC precursor	*Rickettsia typhi*	Wilmington	921	RT0216
NP_220612.1	Outer membrane protein TolC precursor	*Rickettsia prowazekii*	Madrid E	827	RP224
ZP_04699593.1	Type I secretion outer membrane protein TolC	*Rickettsia prowazekii*	Rp22	850	ADE29736
YP_247001.1	TolC family type I secretion outer membrane protein	*Rickettsia felis*	URRWXCal2	795	RF_0985
YP_001493153.1	TolC family type I secretion outer membrane protein	*Rickettsia akari*	Hartford	792	A1C_01670
ZP_04699596.1	Type I secretion outer membrane protein TolC	*Rickettsia endosymbiont of Ixodes scapularis*		757	ZP_04699596
YP_001499115.1	Type I secretion outer membrane protein TolC	*Rickettsia massiliae*	MTU5	751	RMA_0312
ZP_00142231.1	Outer membrane protein tolC precursor	*Rickettsia sibirica*	246	749	Rsib_orf.398
NP_359943.1	Outer membrane protein tolC precursor	*Rickettsia conorii*	Malish 7	748	RC0306
YP_002844973.1	Type I secretion outer membrane protein	*Rickettsia africae*	ESF-5	746	RAF_ORF0284

	TolC				
YP_004764024.1	Outer membrane protein tolC	*Rickettsia heilongjiangensis*	054	745	Rh054_06640
YP_002916267.1	Type I secretion outer membrane protein, TolC precursor	*Rickettsia peacockii*	Rustic	744	RPR_01710
YP_001494423.1	Outer membrane protein tolC precursor	*Rickettsia rickettsii*	Sheila Smith	743	AIG_01745
YP_001649672.1	Type I secretion outer membrane protein	*Rickettsia rickettsii*	Iowa	743	RrIowa_0370
YP_001491990.1	Outer membrane protein tolC precursor	*Rickettsia canadensis*	McKiel	720	AIE_01300
YP_537404.1	Type I secretion outer membrane protein TolC	*Rickettsia bellii*	RML369-C	659	RBE_0234
YP_001496685.1	Type I secretion outer membrane protein TolC	*Rickettsia bellii*	OSU 85-389	657	A1l_06660

Table S4. Primary and secondary antibodies used in the study.

Antibody target	**Host animal**	**Dilution**	**Antibody Source**
RT0218	Rabbit (polyclonal)	1:500	This study
EF-Ts (*R. typhi*)	Rabbit (polyclonal)	1:2000	This laboratory (8)
GAPDH	Mouse (monoclonal)	1:1000	Abcam
R. typhi	Rat (serum)	1:500	This laboratory
Mouse IgG, HRP-linked	Sheep	1:5000	GE Healthcare
Rabbit IgG, HRP-linked, F (ab')$_2$ fragment	Donkey	1:5000	GE Healthcare
Rabbit IgG, Alexa Fluor 594-conjugated	Goat	1:300	Molecular Probes
Rat IgG, Alexa Fluor 488-conjugated	Goat	1:200	Molecular Probes

Table S5. Peptides identified in a unique immogenic band (approximately 80kDa) from cellular pellets and culture supernatant *E. coli* cells by mass spectrometry.

Sample	**Identified peptides matching RARP-1**	**Position (aa)**	**Protein probability**	**Number of peptides**
Purified RARP-1 from C600-pTrc-*RT0218*	ALLDVYPILQEK	511-522	0.99	11
	SYPVSSSNVTLQK*	386-398	0.99	
	KLVESFIK	466-473	0.99	
	NNLFELR	504-510	0.99	
	FISATNNEMYK	163-173	0.99	
	LTEHATLEQMDMYGYIK	434-450	0.99	
	LFNHDTNLNNVEK	230-242	0.88	
	NSTGETLLTAAIYNDNYYLAK	523-543	0.84	
	TAAIYNDNYYLAK	531-543	0.30	
	LFLPDDDIVLGK*	422-433	0.24	
	AAIYNDNYYLAK	532-543	0.84	
Purified RARP-1 from C600-pTrc-*RT0216-RT0217-RT0218*	SYPVSSSNVTLQK	386-398	0.99	14
	ALLDVYPILQEK	511-522	0.99	
	FISATNNEMYK	163-173	0.99	
	VNNETSELAK	403-412	0.99	
	KIDNDQIINNQAK	373-385	0.99	
	SRPNSLALEK	363-372	0.99	
	LTEHATLEQMDMYGYIK	434-450	0.99	
	KIDNDQIINNQAK	373-385	0.99	
	FISATNNEMYK	163-173	0.99	
	LVESFIK	467-473	0.98	
	NNLFELR	504-510	0.96	
	FLVIR	544-548	0.96	
	EAVLPNDYK	174-182	0.94	
	ESQQIDPR	86-93	0.84	
Concentrated culture supernatant from C600-pTrc-*RT0218*	**SYPVSSSNVTLQK***	386-398	0.99	3
	YDIDINK	474-480	0.70	
	NETQM**LFLPDDDIVLGK***	417-433	0.85	

Deamidated amino acid residues are shown in red.

*The peptides found both in the culture supernatant and purified protein from the cellular pellets are shown in bold.

REFERENCES

1. **Biegert, A., and J. Soding. 2008.** De novo identification of highly diverged protein repeats by probabilistic consistency. Bioinformatics 24:807-814.
2. **Crooks, G. E., G. Hon, J. M. Chandonia, and S. E. Brenner.** 2004. WebLogo: a sequence logo generator. Genome Res 14:1188-1190.
3. **Edgar, R. C.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792-1797.
4. **Gillespie, J. J., A. R. Wattam, S. A. Cammer, J. L. Gabbard, M. P. Shukla, O. Dalay, T. Driscoll, D. Hix, S. P. Mane, C. Mao, E. K. Nordberg, M. Scott, J. R. Schulman, E. E. Snyder, D. E. Sullivan, C. Wang, A. Warren, K. P. Williams, T. Xue, H. S. Yoo, C. Zhang, Y. Zhang, R. Will, R. W. Kenyon, and B. W. Sobral.** 2011. PATRIC: the comprehensive bacterial bioinformatics resource with a focus on human pathogenic species. Infect Immun 79:4286-4298.
5. **Koronakis, V., A. Sharff, E. Koronakis, B. Luisi, and C. Hughes.** 2000. Crystal structure of the bacterial membrane protein TolC central to multidrug efflux and protein export. Nature 405:914-919.
6. **Mosavi, L. K., T. J. Cammett, D. C. Desrosiers, and Z. Y. Peng.** 2004. The ankyrin repeat as molecular architecture for protein recognition. Protein Sci 13:1435-1448.
7. **Petersen, T. N., S. Brunak, G. von Heijne, and H. Nielsen.** 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods 8:785-786.
8. **Rahman, M. S., N. C. Ammerman, K. T. Sears, S. M. Ceraul, and A. F. Azad.** 2010. Functional characterization of a *phospholipase A(2)* homolog from *Rickettsia typhi*. J Bacteriol 192:3294-3303.