Supplementary document S4. Structural modeling of VirB8 family proteins.

Structural alignment for eight proteins of the VirB8 family

For structural modeling, target sequences were computationally aligned to a set of templates and manually curated. The template alignment was constructed with eight structures downloaded from the Protein DataBank (Berman et al. 2000): 4MEI, VirB8 of *Bartonella tribocorum*; 2BHM, VirB8 of *Brucella suis* (Terradot et al. 2005); 4LSO, VirB8 of *Bartonella quintana*; 4JF8, TrwG of *Bartonella birtlesii*; 4KZ1, VirB8 of *Bartonella grahamii*; 2CC3, VirB8 of *Agrobacterium tumefaciens* (Bailey et al. 2006); 4NHF (chain A), TrwG of *Bartonella grahamii*; 4O3V, RvhB8-II of *Rickettsia typhi* (**fig. 1**). (*A*) Structural alignment generated using SPDBV/Deep View (Guex and Peitsch 1997). Sequence coordinates are described in panel B. Magenta cylinders and gray arrows depict the α-helices and β-strands for the VirB8 structure of *Bartonella tribocorum*, with highlighted regions depicting the conserved regions utilized to align target sequences to the templates (see Materials and Methods for further information). (*B*) Protein sequences retrieved from UniProt that correspond to the structures shown in panel *A*. Red coloring depicts the portion of the protein included in each respective structure, with regions shown in the alignment in bold. Yellow highlighting depicts regions from the *B*. *suis* and *R*. *typhi* proteins that are absent from the structures.

Analysis of a putative RvhB8-I/RvhB8-II heterodimer

Examination of a potential heterodimer formed by *R*. *typhi* RvhB8-I (YP_067242, modeled structure) and RvhB8-II (YP_067240, PDBID: 4O3V) proteins (**fig. 2**). See text for explanation of protein modeling procedures. (*A*) Ribbon representation of modeled RvhB8-I/RvhB8-II heterodimer. The major dimerization site involving residues of α-helix α1 and the 'NPxG' motif is boxed and enlarged at right. RvhB8-II residues are denoted with primes. Note: the minor dimerization site involving residues within the loop between α-helices α1 and α2 is not supported in this model. (*B*) Sequence alignment of *R*. *typhi* RvhB8-I and RvhB8-II proteins, with secondary structure assignment. Sequences are the globular domains depicted in panel *A*. For each protein, residues involved in the dimerization interface are within black boxes. Invariant residues are highlighted yellow. Magenta cylinders and gray arrows depict the α-helices and β-strands for the RvhB8-I and RvhB8-II structures (see **fig. 5** of the manuscript). For RvhB8-II, the brown star denotes the break in the structure, with missing residues colored white. Note: residues involved in the minor dimerization site of the RvhB8-I and RvhB8-II homodimers are highlighted green.

Assessing the structures of VirB8 family proteins lacking a conserved 'NPxG' motif

Structural models generated for 16 VirB8 sequences exhibiting divergent 'NPxG' motifs within the major dimerization site (**fig. 3**). See text for explanation of protein modeling procedures. (*A*) N**W**DG of LvhB8, *Legionella longbeachae* D-4968 (EEZ95948); (*B*) N**L**LG of hypothetical protein XPN_1731, *Xanthomonas arboricola* pv. pruni MAFF 301427 (GAE59825); (*C*) N**G**MG of TriG, *Escherichia coli* TW00353 (EKI31743); (*D*) N**D**IG of VirB8, *Pseudomonas putida* LD209 (AHC05756); (*E*) N**I**FG of hypothetical protein F892_00045, *Acinetobacter* sp. NIPH 2168 (ENX24895); (*F*) NPT**D** of competence protein, *Helicobacter pylori* Hp A-17 (EJB46117); (*G*) N**S**LG of VirB8, *Ralstonia solanacearum* CMR15 (CBJ36145); (*H*) N**S**RG of VirB8, *Pseudomonas syringae* pv. syringae SM (EPF64149); (*I*) **R**PTG of VirB8, *Hydrogenophaga* sp. T4 (EWS64160); (*J*) **T**PI**S** of hypothetical protein CPS_0790, *Colwellia psychrerythraea* 34H (AAZ25142); (*K*) NPH**L** of conjugal transfer protein TrbF, *Legionella pneumophila* Corby (YP_001249519); (*L*) N**G**TG of type IV secretion protein, *Burkholderia multivorans* CGD1 (WP_006402582); (*M*) N**A**LG of VirB8, *Xanthomonas citri* subsp. citri (NP_942615); (*N*) NPL**A** of conjugative transfer protein, *Stenotrophomonas maltophilia* EPM1 (WP_005410125); (*P*) N**Y**FG of VirB8, *Yersinia pestis* biovar Microtus str. 91001 (NP_995427); note, this structure is shown in **fig. 6*G*** of the manuscript. Residues in bold define all of these 'NPxG' variants.

REFERENCES

Bailey S, Ward D, Middleton R, Grossmann JG, Zambryski PC. 2006. Agrobacterium tumefaciens VirB8 structure reveals potential protein-protein interaction sites. Proc. Natl. Acad. Sci. U. S. A. 103:2582–2587.

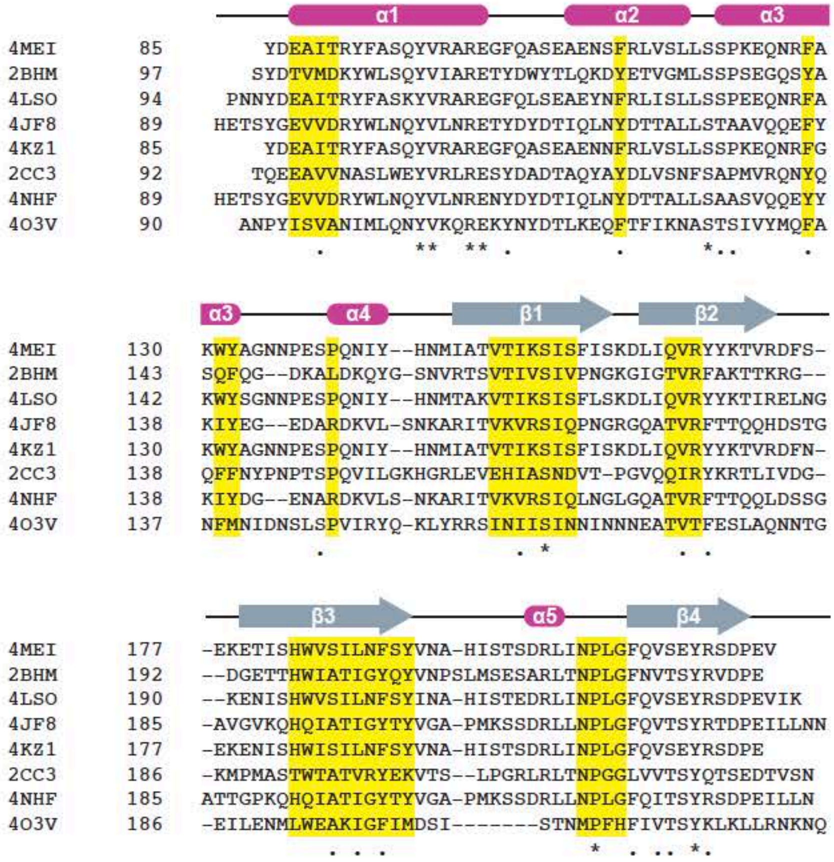
Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H, Shindyalov IN, Bourne PE. 2000. The Protein Data Bank. Nucleic Acids Res. [Internet] 28:235–242. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=102472&tool=pmcentrez&rendertype=abstract

Guex N, Peitsch MC. 1997. SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis [Internet] 18:2714–2723. Available from: http://www.ncbi.nlm.nih.gov/pubmed/9504803

Terradot L, Bayliss R, Oomen C, Leonard GA, Baron C, Waksman G. 2005. Structures of two core subunits of the bacterial type IV secretion system, VirB8 from Brucella suis and ComB10 from Helicobacter pylori. Proc. Natl. Acad. Sci. U. S. A. 102:4596–4601.

Fig. 1 (Document S4)

A



```
4MEI   85   YDEAITRYFASQYVRAREGFQASEAENSFRLVSLLSSPKEQNRFA
2BHM   97  SYDTVMDKYWLSQYVIARETYDWYTLQKDYETVGMLSSPSEGQSYA
4LSO   94 PNNYDEAITRYFASKYVRAREGFQLSEAEYNFRLISLLSSPEEQNRFA
4JF8   89 HETSYGEVVDRYWLNQYVLNRETYDYDTIQLNYDTTALLSTAAVQQEFY
4KZ1   85   YDEAITRYFASQYVRAREGFQASEAENNFRLVSLLSSPKEQNRFG
2CC3   92  TQEEAVVNASLWEYVRLRESYDADTAQYAYDLVSNFSAPMVRQNYQ
4NHF   89 HETSYGEVVDRYWLNQYVLNRENYDYDTIQLNYDTTALLSAASVQQEYY
4O3V   90   ANPYISVANIMLQNYVKQREKYNYDTLKEQFTFIKNASTSIVYMQFA
                .       **  ** .        .       *..     .

4MEI  130 KWYAGNNPESPQNIY--HNMIATVTIKSISFISKDLIQVRYYKTVRDFS-
2BHM  143 SQFQG--DKALDKQYG-SNVRTSVTIVSIVPNGKGIGTVRFAKTTKRG--
4LSO  142 KWYSGNNPESPQNIY--HNMTAKVTIKSISFLSKDLIQVRYYKTIRELNG
4JF8  138 KIYEG--EDARDKVL-SNKARITVKVRSIQPNGRGQATVRFTTQQHDSTG
4KZ1  130 KWYAGNNPESPQNIY--HNMIATVTIKSISFISKDLIQVRYYKTVRDFN-
2CC3  138 QFFNYPNPTSPQVILGKHGRLEVEHIASNDVT-PGVQQIRYKRTLIVDG-
4NHF  138 KIYDG--ENARDKVLS-NKARITVKVRSIQLNGLGQATVRFTTQQLDSSG
4O3V  137 NFMNIDNSLSPVIRYQ-KLYRRSINIISINNINNNEATVTFESLAQNNTG
                .              . *         . .

4MEI  177 -EKETISHWVSILNFSYVNA-HISTSDRLINPLGFQVSEYRSDPEV
2BHM  192 --DGETTHWIATIGYQYVNPSLMSESARLTNPLGFNVTSYRVDPE
4LSO  190 --KENISHWVSILNFSYINA-HISTEDRLINPLGFQVSEYRSDPEVIK
4JF8  185 -AVGVKQHQIATIGYTYVGA-PMKSSDRLLNPLGFQVTSYRTDPEILLNN
4KZ1  177 -EKENISHWISILNFSYVNA-HISTSDRLINPLGFQVSEYRSDPE
2CC3  186 -KMPMASTWTATVRYEKVTS--LPGRLRLTNPGGLVVTSYQTSEDTVSN
4NHF  185 ATTGPKQHQIATIGYTYVGA-PMKSSDRLLNPLGFQITSYRSDPEILLN
4O3V  186 -EILENMLWEAKIGFIMDSI-------STNMPFHFIVTSYKLKLLRNKNQ
                  . . .                 *  . .. *.
```

B

>tr|A9IWN6|A9IWN6_BART1 VirB8 protein OS=Bartonella tribocorum (strain CIP 105476 / IBS 506)
MKSDALEEYIKEARSFDIDRMHGMRVRMKISMALTVLFGLMTIALALAVAALTPLKTVEPFVIRVDNSTGIIDTVSALKESPSD**YDEAITRYFASQYVRARE**
GFQASEAENSFRLVSLLSSPKEQNRFAKWYAGNNPESPQNIYHNMIATVTIKSISFISKDLIQVRYYKTVRDFSEKETISHWVSILNFSYVNAHISTSDRLI
NPLGFQVSEYRSDPEVIK

>sp|Q7CEG3|VIRB8_BRUSU Type IV secretion system protein virB8 OS=Brucella suis biovar 1 (strain 1330)
MFGRKQSPQKSVKNGQGNAPSVYDEALNWEAAHVRLVEKSERRAWKIAGAFGTITVLLGIGIAGMLPLKQHVPYLVRVNAQTGAPDILTSLDEKSV**SYDTVM**
DKYWLSQYVIARETYDWYTLQKDYETVGMLSSPSEGQSYASQFQGDKALDKQYGSNVRTSVTIVSIVPNGKGIGTVRFAKTTKRTNETGDGETTHWIATIGY
QYVNPSLMSESARLTNPLGFNVTSYRVDPEMGVVQ

>sp|Q6FYW3|VIRB8_BARQU Type IV secretion system protein virB8 OS=Bartonella quintana (strain Toulouse)
MKNSLIKIRKSLVKSDAFDEYVKEARSFDIDRMHSLQQRMRIAMTLTVLFGLMTIALALAVAALTPLKTVEPFVIRVDNSTGIIETVSALKET**PNNYDEAIT**
RYFAGKYVRAREGFQLSEAEYNFRLISLLSSPEEQNRFAKWYSGNNPESPQNIYHNMTAKVTIKSISFLSKDLIQVRYYKTIRELNGKENISHWVSILNFSY
INAHISTEDRLINPLGFQVSEYRSDPEVIK

>tr|D0AAZ5|D0AAZ5_9RHIZ TrwG component of type IV secretion system OS=Bartonella birtlesii
MKKQQAKQVTAEKLKSYYEESRGLERELINEFVRSRRTAWRVANVVGIFGLFGMICGVVGFSQPAPTPLVLRVDNTTGAVDVISVMRE**HETSYGEVVDRYWL**
NQYVLNRETYDYDTIQLNYDTTALLSTAAVQQEFYKIYEGEDARDKVLSNKARITVKVRSIQPNGRGQATVRFTTQQHDSTGAVGVKQHQIATIGYTYVGAP
MKSSDRLLNPLGFQVTSYRTDPEILLNN

>tr|C6AER9|C6AER9_BARGA VirB8 protein OS=Bartonella grahamii (strain as4aup)
MKSDALEEYIKEARSFDIDRMHSMRVRMKISMALTVLFGLMTIALALAVAALTPLKTVEPFVIRVDNSTGIIDTVSALKESPND**YDEAITRYFASQYVRARE**
GFQASEAENNFRLVSLLSSPKEQNRFGKWYAGNNPESPQNIYHNMIATVTIKSISFISKDLIQVRYYKTVRDFNEKENISHWISILNFSYVNAHISTSDRLI
NPLGFQVSEYRSDPEVIK

>sp|P17798|VIRB8_AGRFC Protein virB8 OS=Agrobacterium fabrum (strain C58 / ATCC 33970)
MKGSEYALLVARETLAEHYKEVEAFQTARAKSARRLSKVIAAVATIAVLGNVAQAFTIATMVPLIRLVPVYLWIRPDGTVDSEVSVSRLPA**TQEEAVVNASL**
WEYVRLRESYDADTAQYAYDLVSNFSAPMVRQNYQQFFNYPNPTSPQVILGKHGRLEVEHIASNDVTPGVQQIRYKRTLIVDGKMPMASTWTATVRYEKVTS
LPGRLRLTNPGGLVVTSYQTSEDTVSNAGHSEP

>tr|C6AAT5|C6AAT5_BARGA TrwG protein OS=Bartonella grahamii (strain as4aup)
MKTKQVKPVKAEQLSSYYEESRGLERDLINEFIRSRRTAWRVASTVGLFGLFGMICGVVGFSQPAPTPLVLRVDNTTGAVDVISVMRE**HETSYGEVVDRYWL**
NQYVLNRENYDYDTIQLNYDTTALLSAASVQQEYYKIYDGENARDKVLSNKARITVKVRSIQLNGLGQATVRFTTQQLDSSGATTGPKQHQIATIGYTYVGA
PMKSSDRLLNPLGFQITSYRSDPEILLND

>tr|Q68X84|Q68X84_RICTY VirB8-like protein OS=Rickettsia typhi (strain ATCC VR-144 / Wilmington)
MDQLTNAIKEYIKSGEYFIDARKWYNFKYILPLSHRSFLLLTCTIFTLLLTLIGVNINILLPIQKKVGYLIKDDSEKQATITNTKYSTL**ANPYISVANIMLQ**
NYVKQREKYNYDTLKEQFTFIKNASTSIVYMQFANFMNIDNSLSPVIRYQKLYRRSINIISINNINNNEATVTFESLAQNNTGEILENMLWEAKIGFIMDSI
STSTLHNMPFHFIVTSYKLKLLRNKNQQ

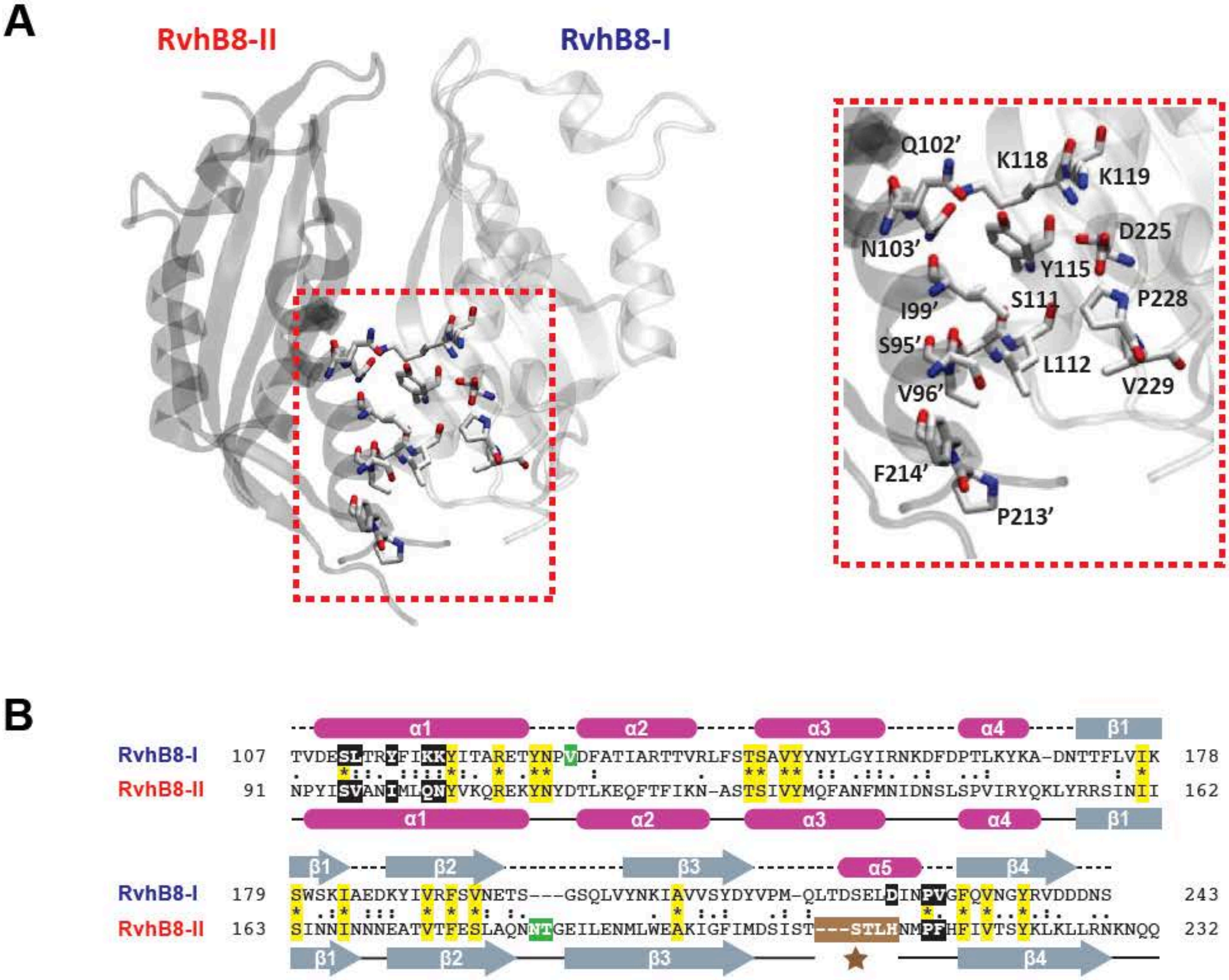


Fig. 2 (Document S4)

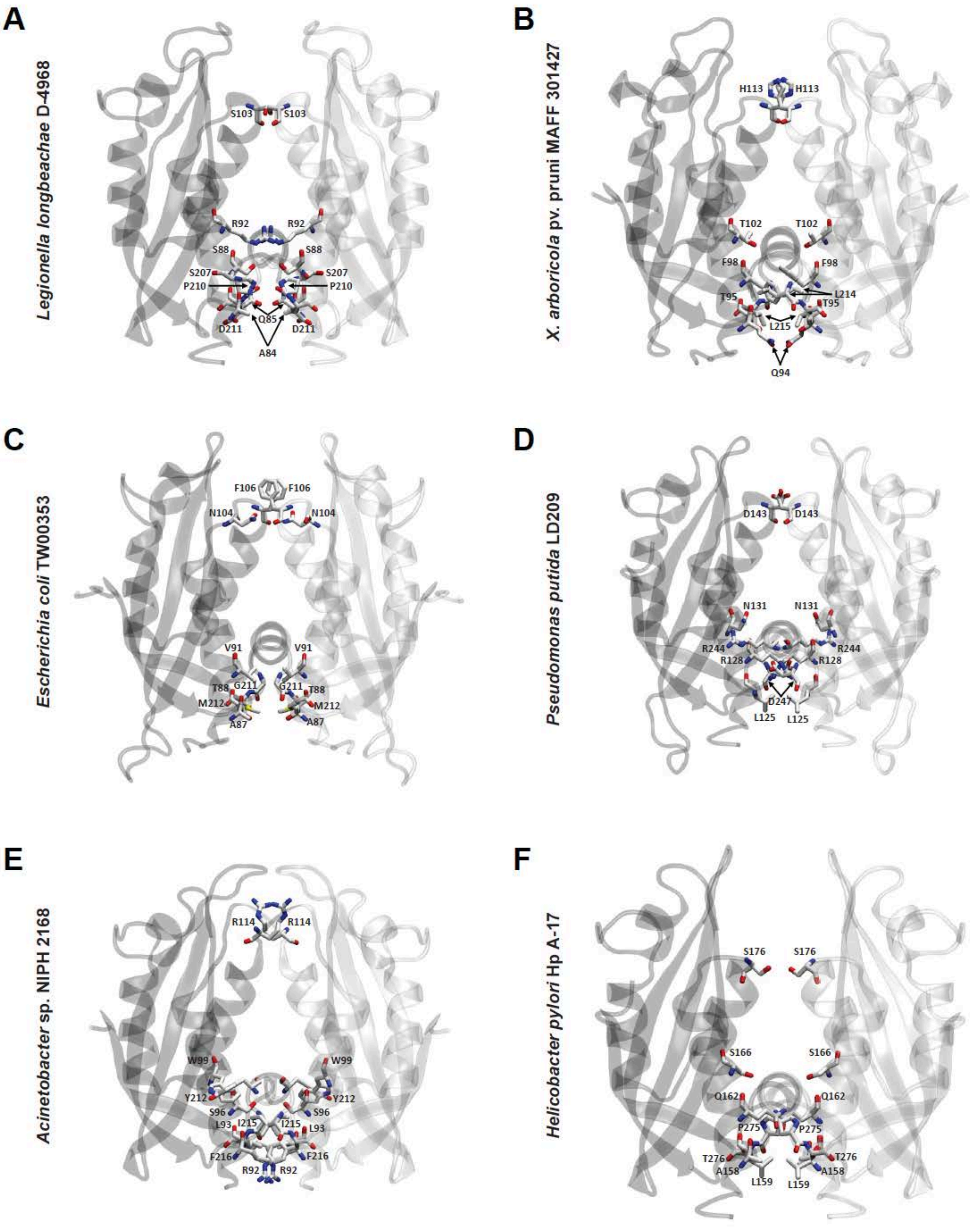


Fig. 3 (Document S4)

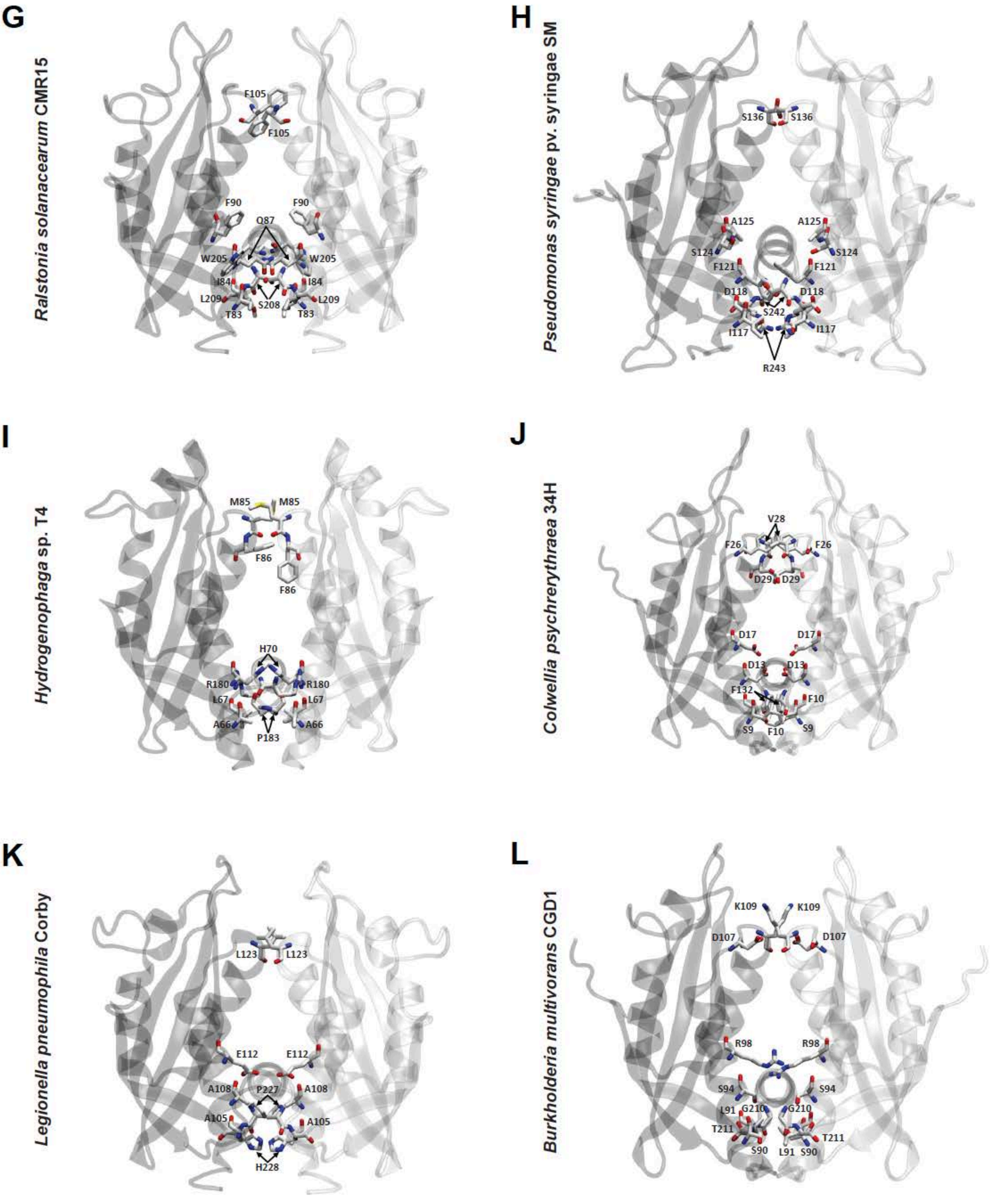


Fig. 3 (Document S4)

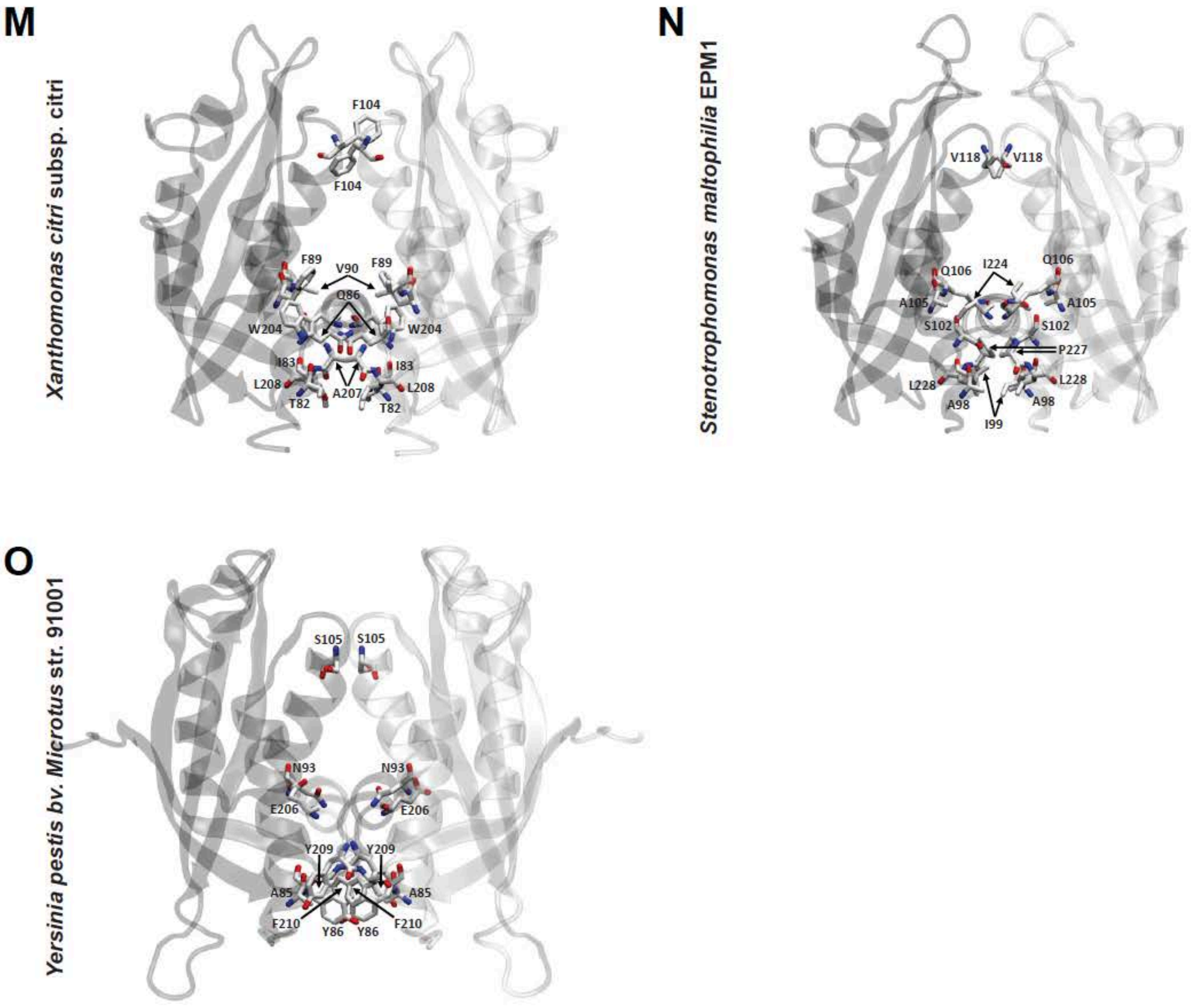


Fig. 3 (Document S4)