

Table S1. Strains, plasmids and oligonucleotides used in this project

Strains		
Name	Genotype	Source/reference
<i>Streptococcus thermophilus</i>	LMG18311 ICESt3	Laboratory collection
DH5 α	<i>F</i> -, Δ (<i>argF-lacZ</i>)U169, <i>phoA</i> , <i>supE44</i> , Δ (<i>lacZ</i>)M15, <i>relA</i> , <i>endA</i> , <i>thi</i> , <i>hsdR</i>	New England Biolabs
BL21 (DE3)	<i>F</i> - <i>ompT gal dcm lon hsdSB(rB- mB-) λ(DE3)</i>	Invitrogen
Plasmids		
Name	Description	Source/reference
pET28a-TRX	6His-tag-TEV protease sequences inserted into pET28a+	Laboratory collection
pET28-TRX-OrfG ₆₄₋₃₃₁	<i>orfG</i> ₆₄₋₃₃₁ cloned in frame of 6His-TEV in pET28-TRX	This project
pET28-TRX-OrfG ₆₄₋₂₀₄	<i>orfG</i> ₆₄₋₂₀₄ cloned in frame of 6His-TEV in pET28-TRX	This project
pET28-TRX-OrfG ₆₄₋₂₁₅	<i>orfG</i> ₆₄₋₂₁₅ cloned in frame of 6His-TEV in pET28-TRX	This project
pET28-TRX-OrfG ₂₂₃₋₃₃₁	<i>orfG</i> ₂₂₃₋₃₃₁ cloned in frame of 6His-TEV in pET28-TRX	This project

Oligonucleotides	
For cloning in pET28-TRX*	
OrfG-BamHI-F	CCCCGGATCCGAGAACCTGTACTTCCAATCATCAAATAAAGTGGGAACACTAC
OrfG-HindIII-R	CCCCAAGCTTCTAATCTTTATCGTTTGCGTAG
OrfG _c -BamHI-F	CCCCGGATCCGAGAACCTGTACTTCCAATCAGCCAACGACCACGTCTCAG
*Restrictions sites are underlined and TEV cleavage encoding sequence are highlighted in italics	
Cloning by site directed mutagenesis (Stop codon insertion by PCR)*	
OrfG ₆₄₋₂₀₄ -Stop-F	ACTTCCTTGGTTCTCAGCGTAAGAGTCTTACAAGCTGGATAC
OrfG ₆₄₋₂₀₄ -Stop-R	GTATCCAGCTTGTGAAGACTCTTACGCTGAGAACCAAGGAAGT
OrfG ₆₄₋₂₁₅ -Stop-F	AGCTGGATACTTTAGCGAATAGGATCAACTCCAACCTCACAG
OrfG ₆₄₋₂₁₅ -Stop-R	CTGTGAGTTGGAGTTGATCCTATTCGCTAAAGTATCCAGCT
* Mutagenized codons are underlined	

Tables S2. Pairwise alignment lengths, identities and similarities between the 13 Gram-negative VirB8 and 4 Gram-positive VirB8-like proteins considered in this comparison (**Table S3a**). The sequences are derived from the structures, i.e. there might be missing residues (see protein description below) corresponding to unobserved regions in the electron density. Values are calculated from the alignment of the sequences corresponding to the overall superposition of the 17 structures, as performed by mTM-align (Dong *et al.*, 2018a). The TM-scores (Dong *et al.*, 2018b) and the pairwise rmsd (Å) calculated from the same overall superposition are also given (**Tables S3b and S3c**). Clusters that appear in the structural tree are coloured as in Fig.5 and numbered for their convenient description in the main text.

For each following PDB code, the source organism, the Uniprot code, the sequence that was used for the experiment (X-Ray or NMR), where appropriate the unobserved residues (which are therefore missing in the sequence) and the resulting coverage of the whole sequence are given hereafter.

- 2cc3: VirB8 from *Agrobacterium tumefaciens*; Uniprot: P17798; Crystallized sequence: 92-237; C-ter (232-237) unobserved in the electron density; Coverage 59%.
- 3ub1: TcpC from *Clostridium perfringens*; Uniprot: Q1PLH8; Crystallized sequence: 99-309, central domain as defined by authors: 104-231; loop β 2- β 3 (190-192) unobserved in the electron density (coverage 35%).
- 3wz3: TraM from Plasmid R64; Uniprot Q9R2H2; Crystallized sequence: 91-230; C-ter (227-230) unobserved in the electron density (coverage 59%).
- 3wz4: IcmL/DotI from *Legionella pneumophila*; Uniprot: O54626; Crystallized sequence: 73-212; C-ter (209-212) unobserved in the electron density (coverage 64%).
- 4akz: VirB8 from *Brucella suis*; Uniprot: Q7CEG3; Crystallized sequence: 97-234 (coverage 57%).
- 4ec6: TraM from *Enterococcus faecalis*; Uniprot: Q8L1C7; Crystallized sequence: 190-322; N-ter (190-213) unobserved in the electron density (coverage 34%).
- 4jf8: TrwG from *Bartonella birtlesii*; Uniprot: D0AAZ5; Crystallized sequence: 63-232; N-ter (63-88) unobserved in the electron density (coverage 62%).
- 4kz1: VirB8 from *Bartonella grahamii*; Uniprot: C6AER9; Crystallized sequence: 51-222; N-ter (51-84) and C-ter (220-222) unobserved in the electron density; Coverage 61%.
- 4lso: VirB8 from *Bartonella quintana*; Uniprot: Q6FYW3; Crystallized sequence: 65-234; N-ter (65-93) unobserved in the electron density; Coverage 60%.
- 4mei: VirB8 from *Bartonella tribocurum*; Uniprot: A9IWN6; Crystallized sequence: 51-222; N-ter (51-84) and C-ter (221-222) unobserved in the electron density; Coverage 61%.
- 4nhf: VirB8 from *Bartonella grahamii*; Uniprot: C6AAT5; Crystallized sequence: 63-233; N-ter (63-88) and C-ter (233) unobserved in the electron density (coverage 61%).
- 4o3v: VirB8 from *Rickettsia typhi*; Uniprot: Q68X84; Crystallized sequence: 61-232; N-ter (61-89), loop β 3- β 4 (208-209) and C-ter (232) unobserved in the electron density (coverage 60%).
- 5aiw: TraH/Orf8 from *Enterococcus faecalis*; Uniprot: Q9AL12; NMR sequence: 56-183 (coverage 69%).
- 5cnl: IcmL-like from *Legionella pneumophila*; Uniprot: Q5ZZ91; Crystallized sequence: 22-176; C-ter (162-176) unobserved in the electron density (coverage 79%).
- 5i97: TraE from *Escherichia coli*; Uniprot: Q17U16; Crystallized sequence: 70-232; N-ter (70-87) and C-ter (231-232) unobserved in the electron density (coverage 59%).
- 6iqt: CagV from *Helicobacter pylori*; Uniprot: O25265; Crystallized sequence: 74-229; N-ter (74-85), loop β 3- β 4 (202-209) and C-ter (229) unobserved in the electron density (coverage 51%).
- 6zgn: OrfG from *Streptococcus thermophilus*; Uniprot: Q70CA4; Crystallized sequence: 65-204; N-ter (65-91), loop before β 1 (134-138) and loop β 2- β 3 (169-175) unobserved in the electron density (coverage 31%).

Table S2a: Sequence comparison

		Gram-negative VirB8												Gram-positive VirB8-like						
		Class I ⁻				Class II ⁻								Class III ⁻				Class I ⁺		
		5i97	4akz	4nhf	4jf8	4mei	4kz1	4lso	2cc3	6iqt	4o3v	3wz4	3wz3	5cnl	6zgn	3ub1	4ec6	5aiw		
5i97	length ident. simil.	143 1 1																		
4akz	length ident. simil.	145 0.414 0.586	138 1 1																	
4nhf	length ident. simil.	149 0.383 0.557	148 0.419 0.534	144 1 1																
4jf8	length ident. simil.	150 0.367 0.573	149 0.422 0.564	145 0.976 0.924	144 1 1															
4mei	length ident. simil.	150 0.200 0.413	145 0.338 0.51	149 0.282 0.456	150 0.287 0.473	136 1 1														
4kz1	length ident. simil.	149 0.215 0.436	144 0.340 0.535	149 0.295 0.456	150 0.300 0.473	136 0.956 0.971	135 1 1													
4lso	length ident. simil.	152 0.204 0.454	150 0.300 0.493	150 0.253 0.460	151 0.265 0.477	141 0.837 0.879	141 0.837 0.894	141 1 1												
2cc3	length ident. simil.	157 0.248 0.331	153 0.242 0.279	158 0.241 0.261	159 0.233 0.346	146 0.240 0.425	146 0.240 0.425	151 0.219 0.431	144 1 1											
6iqt	length ident. simil.	154 0.091 0.227	151 0.139 0.258	157 0.134 0.280	158 0.139 0.279	151 0.133 0.265	150 0.133 0.267	155 0.136 0.252	157 0.115 0.274	135 1 1										
4o3v	length ident. simil.	163 0.147 0.282	158 0.196 0.310	163 0.172 0.319	164 0.195 0.335	152 0.151 0.265	152 0.145 0.322	156 0.160 0.308	156 0.186 0.333	138 1 1										
3wz4	length ident. simil.	165 0.091 0.249	161 0.093 0.248	165 0.091 0.285	166 0.102 0.295	154 0.104 0.292	153 0.085 0.288	158 0.089 0.272	163 0.074 0.282	154 0.104 0.266	162 0.062 0.235	137 1 1								
3wz3	length ident. simil.	166 0.042 0.187	161 0.081 0.205	165 0.097 0.206	166 0.054 0.175	154 0.091 0.260	153 0.084 0.261	159 0.076 0.245	163 0.074 0.227	155 0.058 0.187	161 0.037 0.199	139 0.266 0.453	137 1 1							
5cnl	length ident. simil.	169 0.112 0.231	164 0.110 0.226	167 0.132 0.252	168 0.125 0.262	157 0.096 0.236	157 0.096 0.255	161 0.099 0.248	166 0.060 0.217	158 0.101 0.234	164 0.073 0.220	148 0.264 0.432	149 0.134 0.322	139 1 1						
6zgn	length ident. simil.	152 0.059 0.197	147 0.068 0.204	153 0.059 0.177	153 0.072 0.170	145 0.062 0.172	144 0.063 0.181	150 0.053 0.160	153 0.059 0.163	145 0.076 0.186	149 0.060 0.168	144 0.056 0.188	144 0.069 0.208	146 0.062 0.164	101 1 1					
3ub1	length ident. simil.	169 0.041 0.189	165 0.042 0.188	171 0.035 0.164	171 0.041 0.175	163 0.080 0.178	162 0.086 0.185	168 0.083 0.173	170 0.035 0.171	158 0.101 0.203	165 0.042 0.182	159 0.063 0.182	159 0.063 0.201	160 0.063 0.181	127 0.095 0.323	122 1 1				
4ec6	length ident. simil.	158 0.051 0.190	153 0.046 0.206	157 0.045 0.153	157 0.051 0.159	151 0.046 0.166	150 0.047 0.167	155 0.032 0.155	159 0.038 0.182	147 0.088 0.218	155 0.039 0.194	146 0.055 0.240	146 0.089 0.206	147 0.075 0.197	114 0.105 0.290	129 0.109 0.256	109 1 1			
5aiw	length ident. simil.	165 0.067 0.200	161 0.087 0.193	166 0.108 0.217	167 0.096 0.198	158 0.063 0.203	157 0.064 0.204	162 0.056 0.191	167 0.066 0.186	154 0.065 0.175	161 0.118 0.236	154 0.065 0.214	155 0.071 0.232	159 0.031 0.195	141 0.028 0.177	155 0.077 0.200	144 0.049 0.208	128 1 1		

Table S2b : TM-scores

	Gram-negative VirB8														Gram-positive VirB8-like			
	Class I ⁻				Class II ⁻				Class III ⁻						Class I ⁺			
	5i97	4akz	4nhf	4jf8	4mei	4kz1	4lso	3cc3	6iqt	4o3v	3wz4	3wz3	5cnl	6zgn	3ub1	4ec6	5aiw	
5i97	1																	
4akz	0.9167	1																
4nhf	0.8942	0.9071	1															
4jf8	0.893	0.9259	0.9438	1														
4mei	0.8768	0.8754	0.8789	0.8699	1													
4kz1	0.8854	0.8842	0.879	0.8748	0.9948	1												
4lso	0.8665	0.8619	0.878	0.8702	0.9819	0.984	1											
2cc3	0.825	0.8901	0.8365	0.838	0.9123	0.9154	0.883	1										
6iqt	0.8129	0.8043	0.8053	0.7898	0.8311	0.8298	0.8377	0.8139	1									
4o3v	0.7696	0.7971	0.8158	0.7971	0.8208	0.8247	0.8242	0.8326	0.784	1								
3wz4	0.7218	0.7233	0.7432	0.7402	0.7669	0.7691	0.7674	0.7462	0.7733	0.739	1							
3wz3	0.7054	0.7159	0.7391	0.7371	0.7257	0.7291	0.7221	0.7282	0.7408	0.7414	0.8873	1						
5cnl	0.7108	0.7177	0.7397	0.7324	0.7519	0.75	0.7468	0.7258	0.7738	0.7559	0.8823	0.8142	1					
6zgn	0.733	0.7587	0.7454	0.7442	0.743	0.7445	0.7426	0.7586	0.7462	0.7548	0.7725	0.7626	0.7655	1				
3ub1	0.6146	0.6399	0.6202	0.6237	0.6259	0.6266	0.6342	0.6268	0.6432	0.656	0.6907	0.6641	0.6714	0.8122	1			
4ec6	0.6871	0.7121	0.7052	0.715	0.7119	0.7081	0.7212	0.6703	0.7231	0.6991	0.7363	0.716	0.7169	0.8057	0.7764	1		
5aiw	0.6882	0.6891	0.7054	0.6988	0.6775	0.6757	0.6976	0.6962	0.7257	0.7022	0.7033	0.7047	0.678	0.7036	0.6537	0.656	1	

Table S2c: rmsd (Å)

	Gram-negative VirB8														Gram-positive VirB8-like			
	Class I ⁻				Class II ⁻				Class III ⁻						Class I ⁺			
	5i97	4akz	4nhf	4jf8	4mei	4kz1	4lso	3cc3	6iqt	4o3v	3wz4	3wz3	5cnl	6zgn	3ub1	4ec6	5aiw	
5i97	0																	
4akz	1.426	0																
4nhf	1.628	1.631	0															
4jf8	1.56	1.525	1.316	0														
4mei	2.059	1.837	1.738	1.812	0													
4kz1	2.062	1.818	1.754	1.758	0.326	0												
4lso	2.133	1.86	1.969	2.015	0.602	0.562	0											
2cc3	2.128	1.503	1.996	2.114	1.483	1.416	1.704	0										
6iqt	2.286	2.013	2.256	2.505	1.513	1.395	1.738	1.979	0									
4o3v	2.942	2.172	2.495	2.465	2.061	2	2.348	2.488	2.037	0								
3wz4	3.108	2.843	2.687	2.628	2.096	2.103	2.134	2.456	2.408	2.379	0							
3wz3	3.179	2.95	2.694	2.715	2.799	2.82	2.84	2.759	2.77	2.638	1.691	0						
5cnl	3.045	2.823	2.902	3.152	2.301	2.301	2.475	2.814	2.498	2.452	1.631	2.161	0					
6zgn	2.704	2.444	2.605	2.62	2.401	2.402	2.411	2.314	2.359	2.33	2.262	2.393	2.399	0				
3ub1	3.362	3.271	3.675	3.415	3.312	3.248	3.332	3.263	3.219	3.002	2.849	3.18	3.042	1.99	0			
4ec6	2.841	2.718	2.939	2.767	2.691	2.673	2.558	3.114	2.762	2.757	2.599	2.939	2.97	1.878	2.301	0		
5aiw	2.943	2.777	2.931	2.888	2.774	2.757	2.852	2.997	2.774	3.13	2.939	2.863	3.034	2.474	2.637	3.12	0	

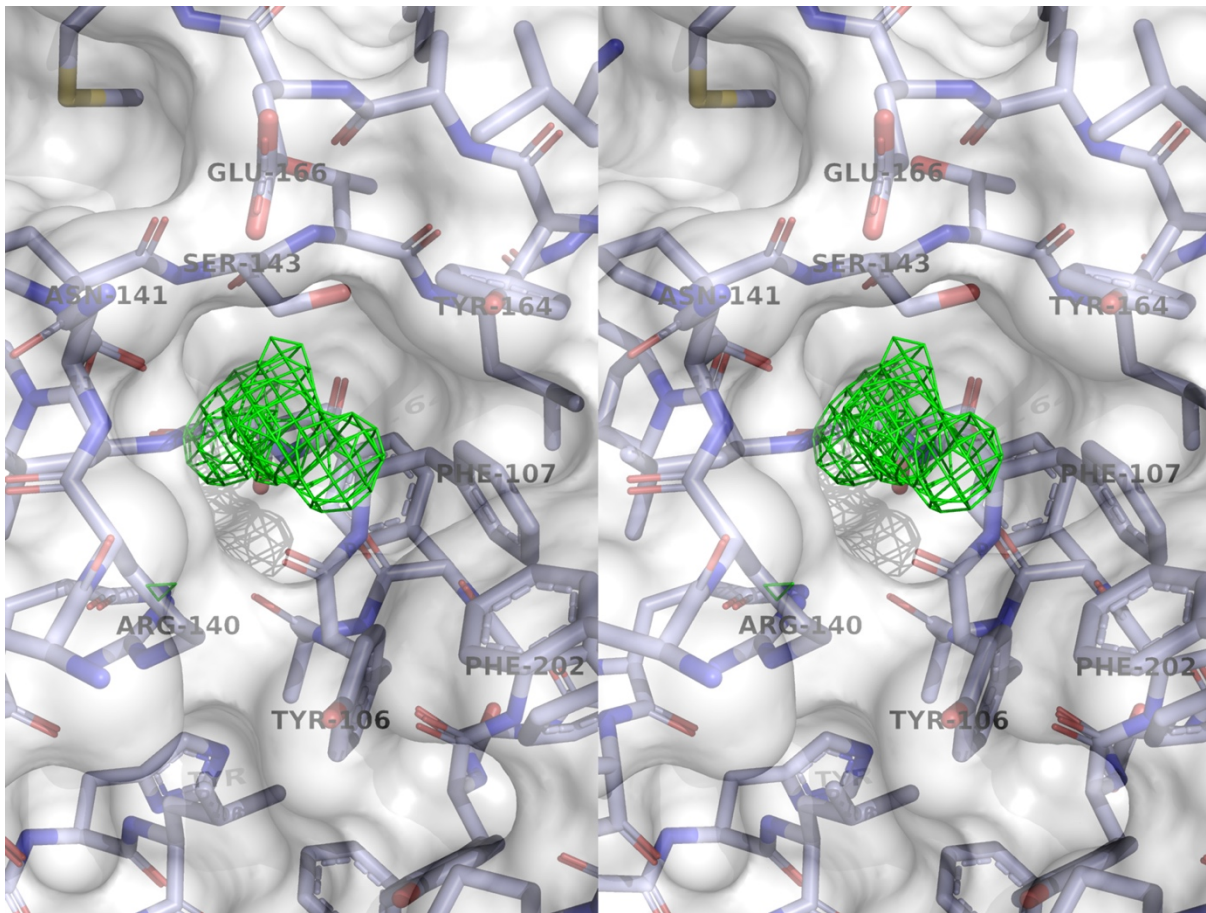


Figure S1: Presence of an unknown ligand bound to OrfG₆₄₋₂₀₄. Stereo view of an unattributed blob of electron density observed in the Fo-Fc map contoured at $+3\sigma$. Surrounding amino acids are shown in sticks and labeled.

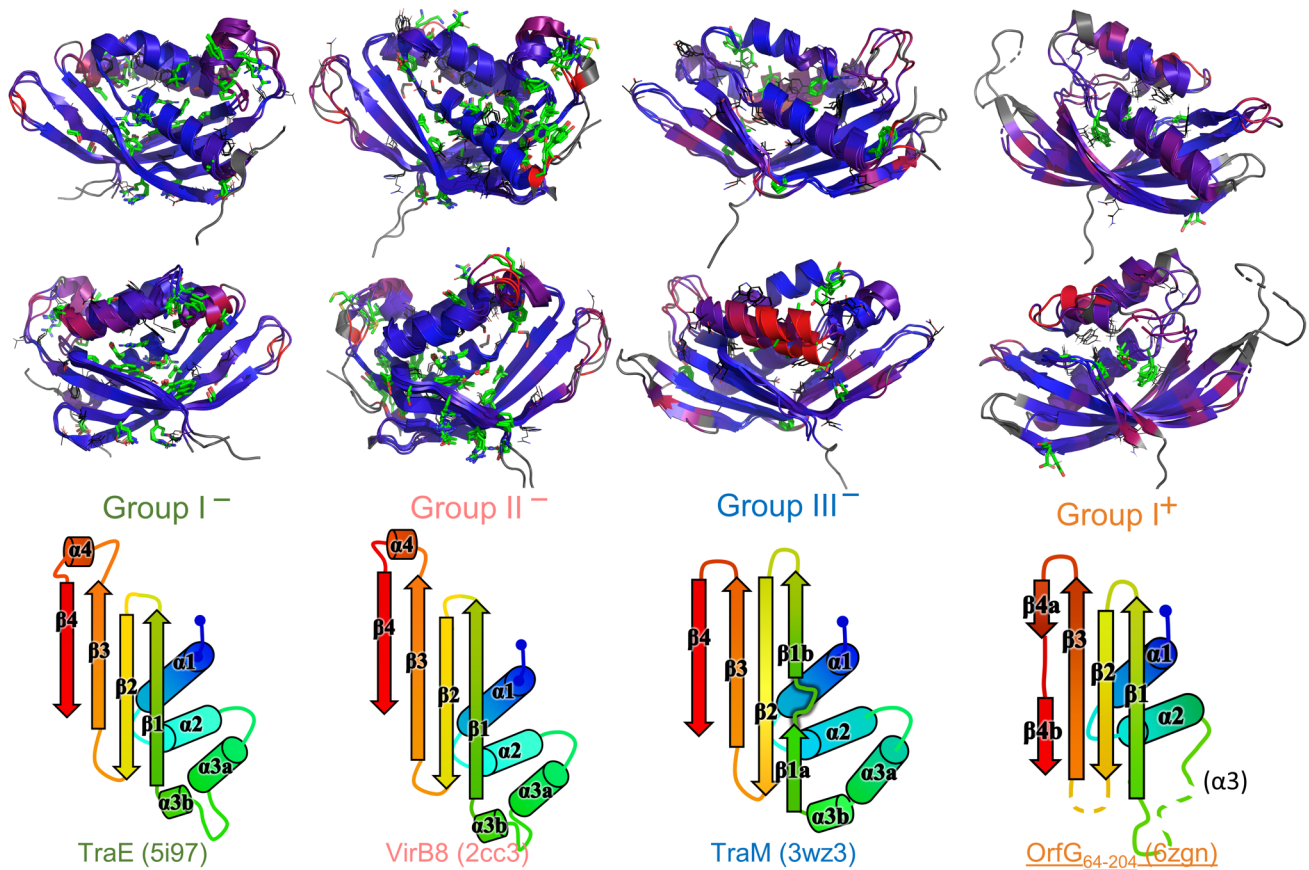


Figure S2: Superposition of VirB8-like structures within each class identified in the structural tree. Structures were superposed by mTM-Align, and represented using the same color code as in Figure 3 (see also Materials and Methods). Amino acid sidechains that are conserved are shown in green sticks, those that share a similar nature within a class are drawn as black lines. For each class, front view, rear view and secondary structure topology of a representative member are shown vertically.

References:

Dong R, Peng Z, Zhang Y & Yang J (2018) mTM-align: an algorithm for fast and accurate multiple protein structure alignment. *Bioinformatics* **34**: 1719-1725.

Dong R, Pan S, Peng Z, Zhang Y & Yang J (2018) mTM-align: a server for fast protein structure database search and multiple protein structure alignment. *Nucleic Acids Res* **46**: W380-W386.