

In the format provided by the authors and unedited.

# Architecture of the type IV coupling protein complex of *Legionella pneumophila*

Mi-Jeong Kwak<sup>1,#</sup>, J Dongun Kim<sup>1,#</sup>, Hyunmin Kim<sup>1</sup>, Cheolhee Kim<sup>2</sup>, James W. Bowman<sup>3</sup>, Seonghoon Kim<sup>1</sup>, Keehyoung Joo<sup>4</sup>, Jooyoung Lee<sup>4</sup>, Kyeong Sik Jin<sup>5</sup>, Yeon-Gil Kim<sup>5</sup>, Nam Ki Lee<sup>2</sup>, Jae U. Jung<sup>3</sup>, Byung-Ha Oh<sup>1,\*</sup>

<sup>1</sup>Department of Biological Sciences, KAIST Institute for the Biocentury, Korea Advanced Institute of Science and Technology, Daejeon 305-701, Korea

<sup>2</sup>Department of Physics, Pohang University of Science and Technology, Pohang, Kyungbuk 790-784, Korea

<sup>3</sup>Department of Molecular Microbiology and Immunology, Keck School of Medicine, University of Southern California, Los Angeles, California, USA

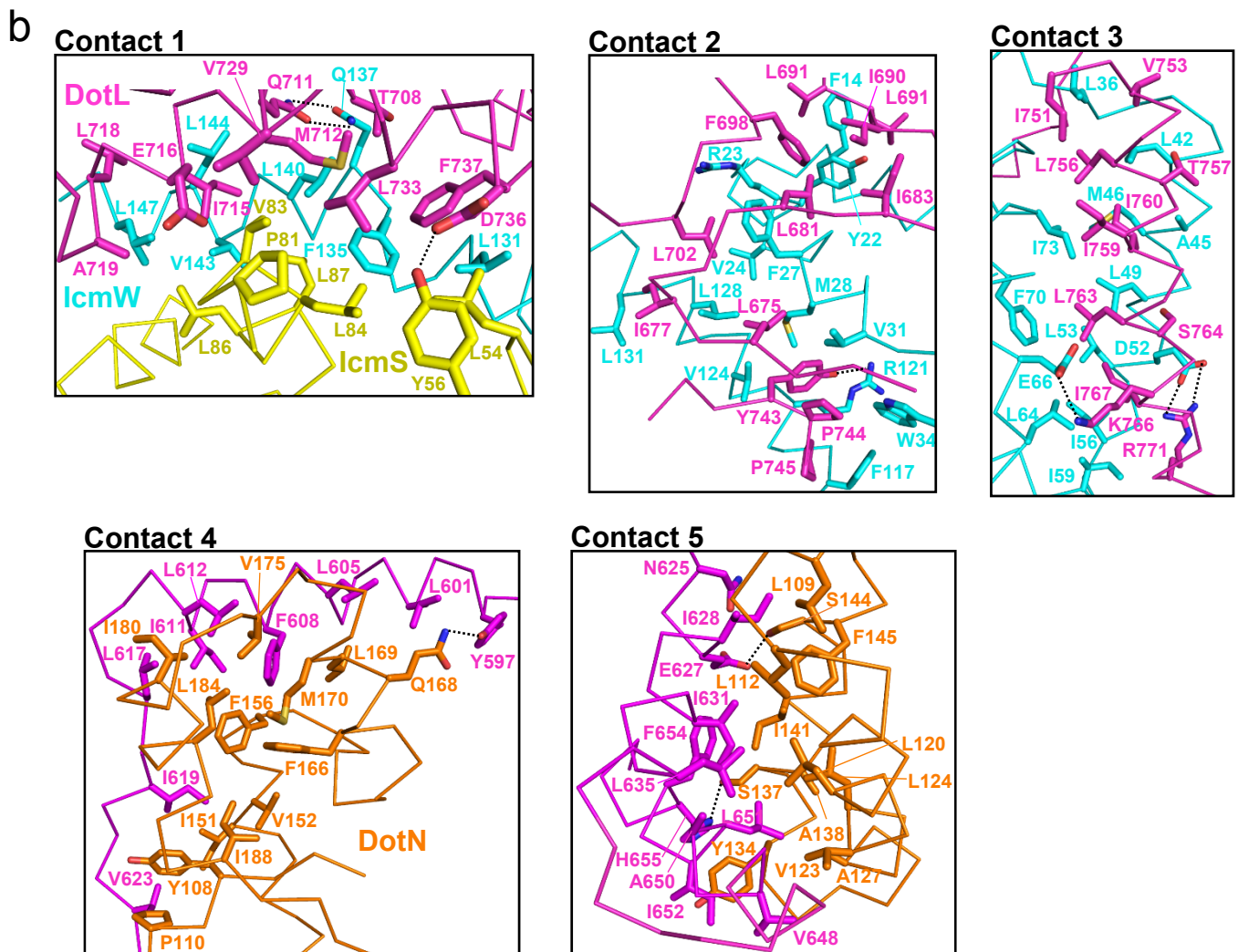
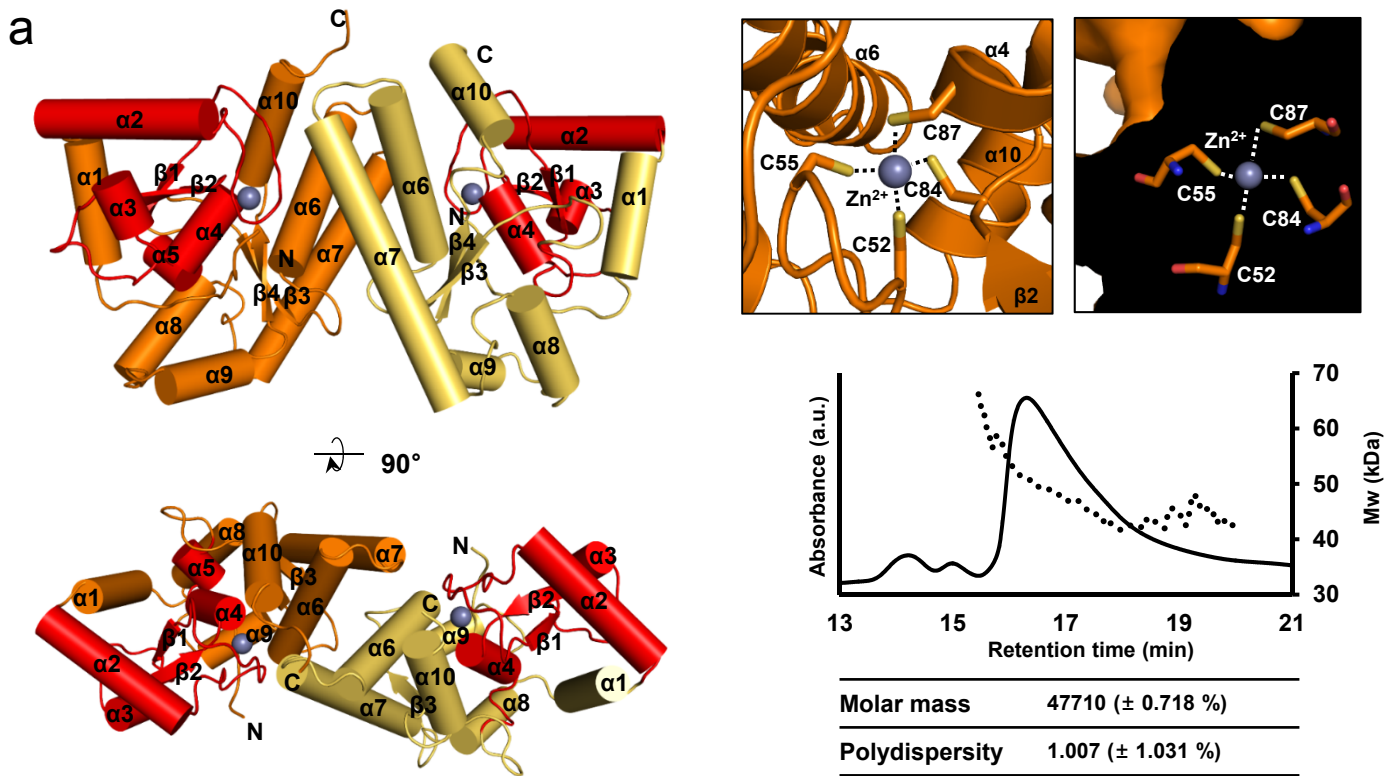
<sup>4</sup>Center for in Silico Protein Science, Korea Institute for Advanced Study, Seoul 130-722, Korea

<sup>5</sup>Pohang Accelerator Laboratory, Pohang University of Science and Technology, Pohang, Kyungbuk, 790-784, Korea

#Co-first authors; \*Corresponding author (e-mail: [bhoh@kaist.ac.kr](mailto:bhoh@kaist.ac.kr))

## Supplementary information

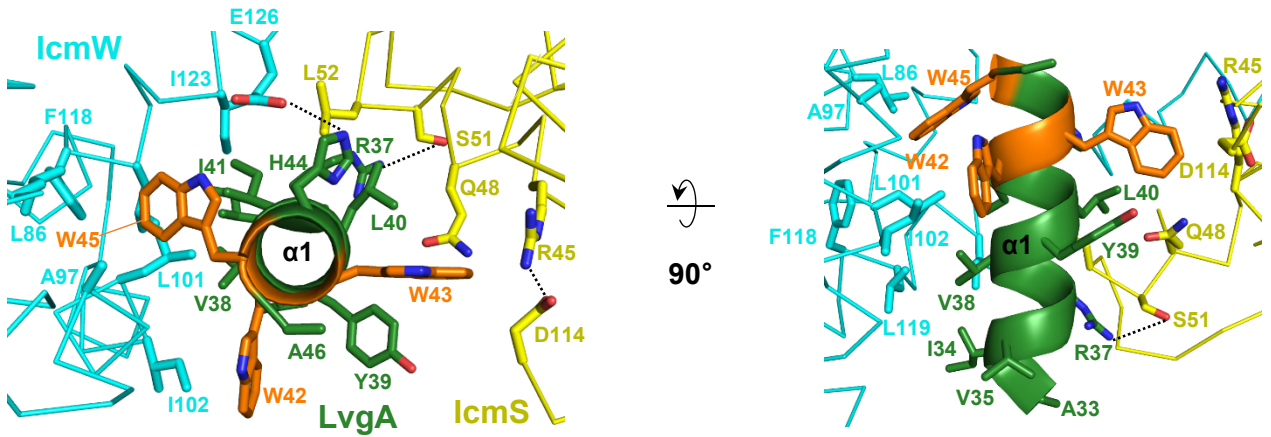
This file contains Supplementary Figures 1-7 and Supplementary Table 1.



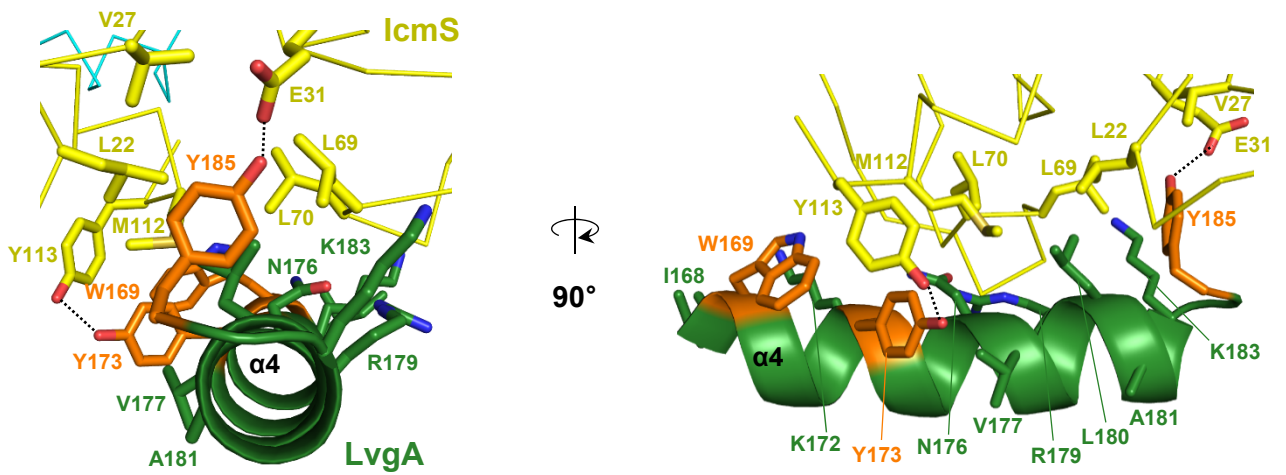
**Supplementary Figure 1. Crystal structure of free DotN and intersubunit interactions of DotL(590-783) with IcmSW and DotN**

(a) Structure of free DotN. (*Left*) Two perpendicular views are shown with the two monomers in different colors and the bound Zn<sup>2+</sup> ions shown as gray spheres. Two antiparallel  $\alpha$ -helices ( $\alpha_6$ ,  $\alpha_7$ ) of one monomer are stacked onto the equivalent  $\alpha$ -helices of the other monomer, as if they form a four-helical bundle. The 62-residue segment structurally similar to the HNH superfamily nucleases is indicated by red color. (*Right, Top*) Two views of the zinc cage. Shown are a ribbon drawing with the four Zn<sup>2+</sup>-coordinating cysteines in sticks and a cut-away view demonstrating the complete burial of Zn<sup>2+</sup>. (*Right, Bottom*) Molecular mass analysis of free DotN by size-exclusion chromatography coupled with multi-angle light scattering analysis.

(b) Detailed intersubunit interactions. Residue-residue contacts are shown for the five interfaces marked in Figure 1a,b. Hydrogen bonds are indicated by dashed lines.

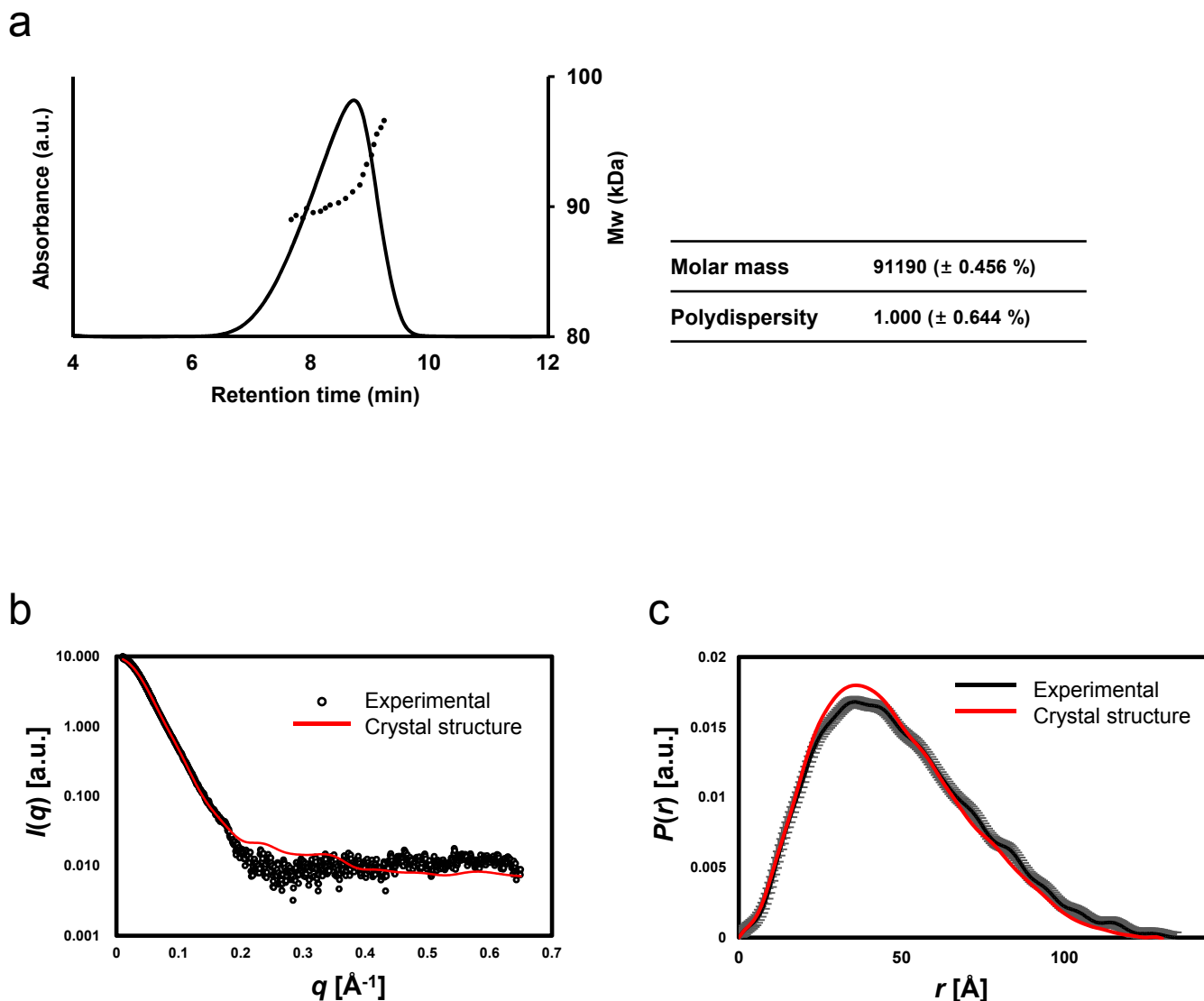


Intermolecular interaction of  $\alpha 1$  of LvgA



Intermolecular interaction of  $\alpha 4$  of LvgA

**Supplementary Figure 2. Highlighted hydrophobic interactions between LvgA and IcmSW**  
 Intersubunit residue-residue contacts are shown. Hydrogen bonds are indicated by dashed lines.



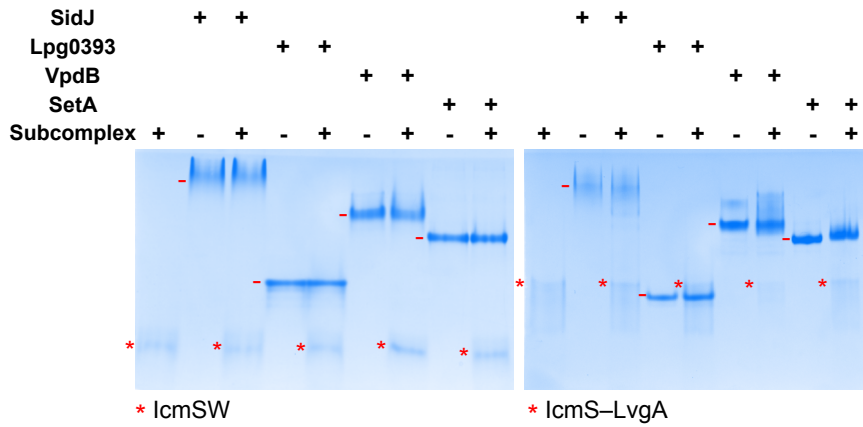
### Supplementary Figure 3. SAXS analysis of DotL(590-783)–DotN–IcmSW–LvgA

(a) Molecular mass analysis of the complex by AF4-MALS.

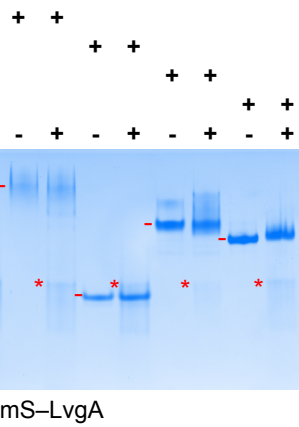
(b) The SAXS curves showing the experimental and the calculated X-ray scattering. The plot shows the scattering intensity  $I(q)$  as a function of  $q$  ( $q = 4\pi\sin\theta/\lambda$ , where  $2\theta$  is the scattering angle and  $\lambda$  is the wavelength). The scattering data were extrapolated to zero concentration and normalized by zero angle scattering intensity  $I(0)$ . The experimental scattering intensities are the average of six successive frames of 5 to 10 s exposure, indicating no sign of radiation damage.

(c) The  $P(r)$  functions showing the experimental and the calculated distance distributions. Distribution of inter-atomic distances,  $P(r)$ , is plotted as a function of distance ( $r$ ).

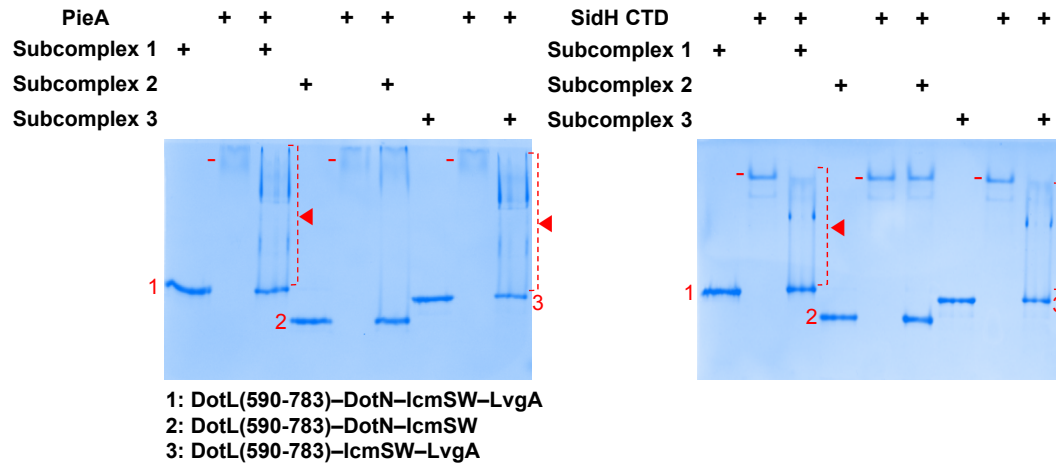
**a**



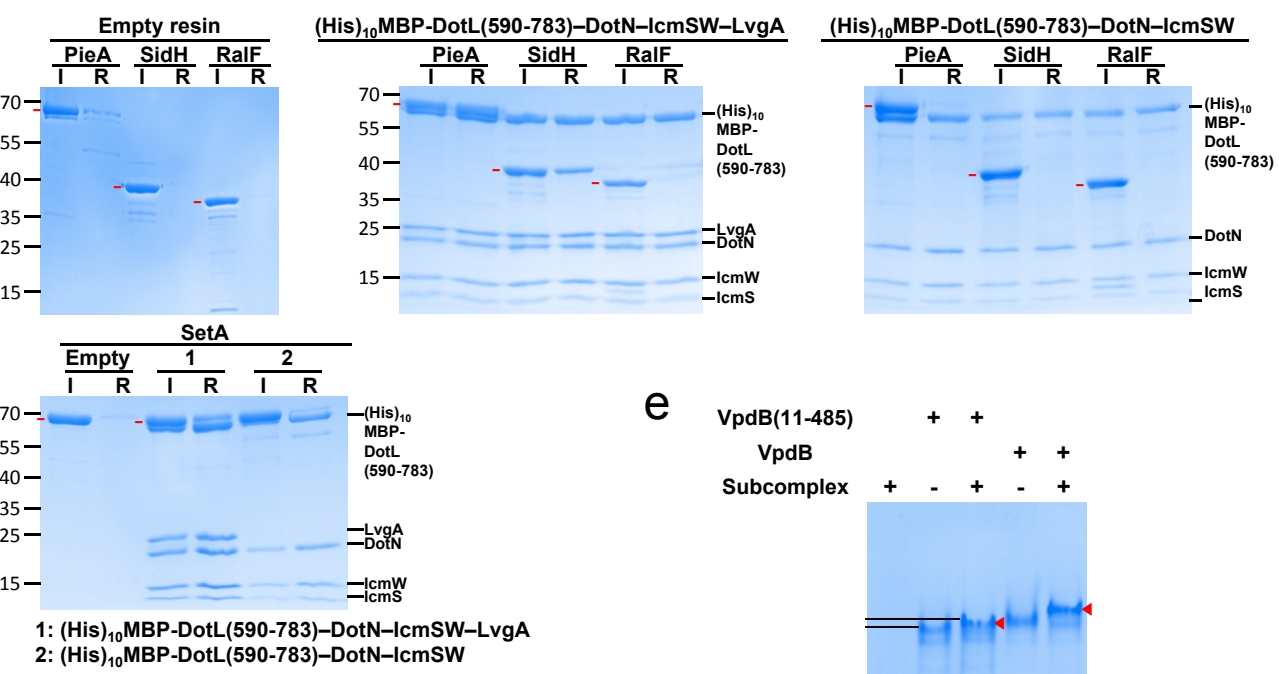
**b**



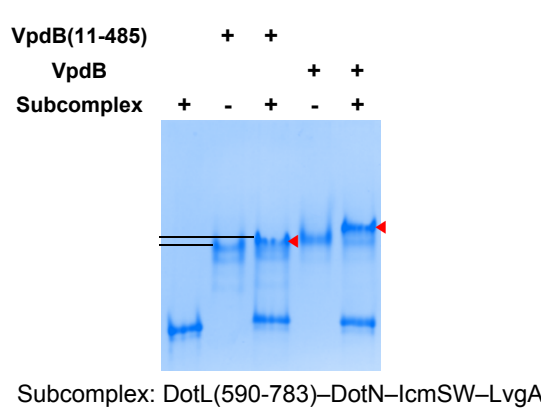
**c**



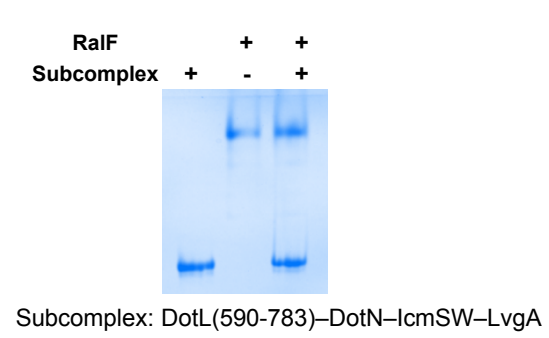
**d**



**e**



**f**

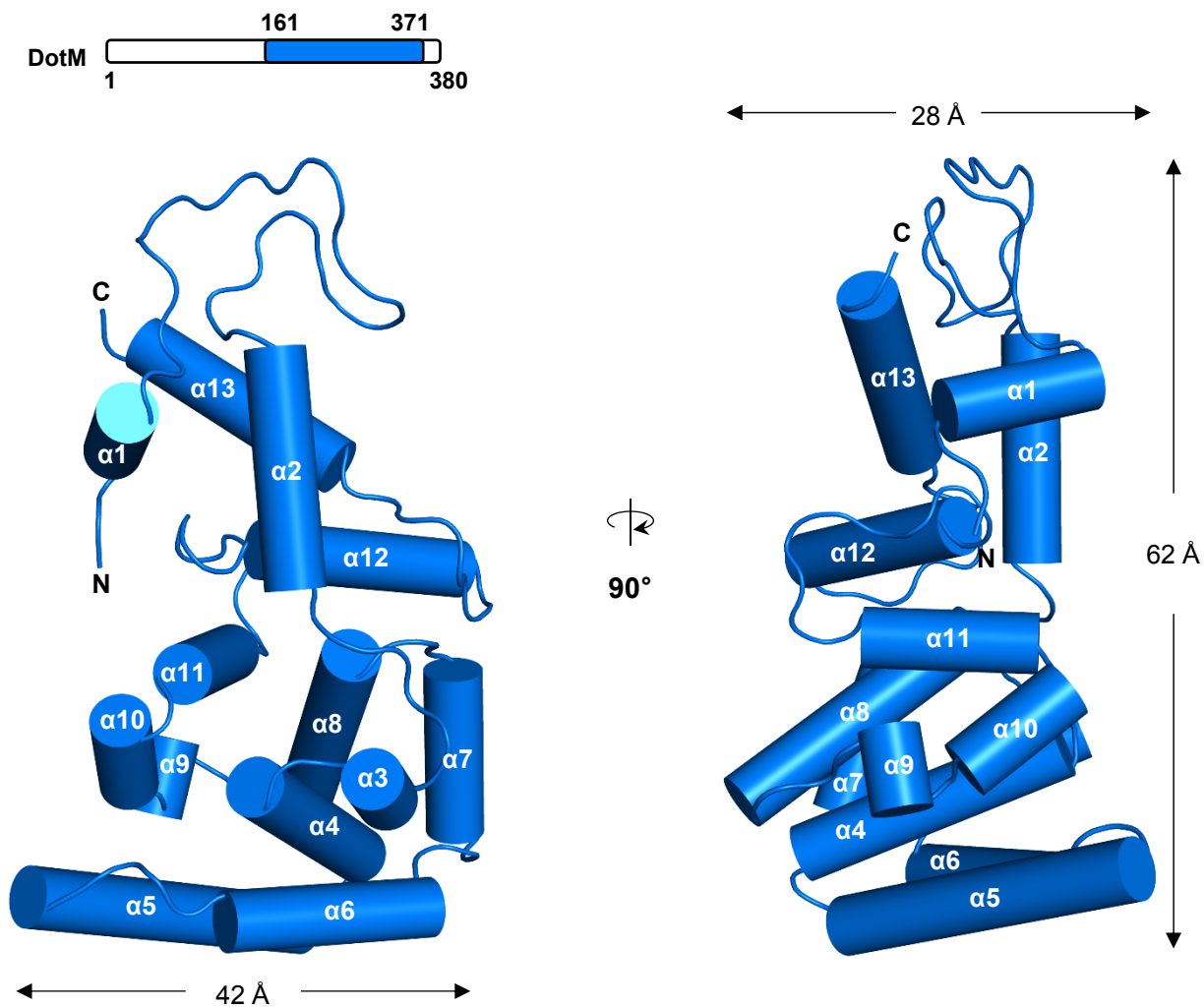


#### Supplementary Figure 4. Protein binding assays

- (a-b) Interaction of IcmSW and IcmS–LvgA with effector proteins. The four indicated effector proteins (6  $\mu$ M) were incubated with IcmSW or IcmS–LvgA at 1:1 molar ratio, and the mixtures were analyzed by native PAGE. (a) The migration of the four effectors did not change upon addition of IcmSW. (b) Tailing of the VpdB or SetA protein band was observed upon addition of IcmS–LvgA.
- (c) Native PAGE analysis for PieA and SidH(1830-2225). The two proteins (6  $\mu$ M) were incubated with the indicated subcomplexes at 1:1 molar ratio. The effector proteins are indicated by '-'. Triangles indicate newly formed protein bands, which are smeared for both effectors. The reason for this migration behavior is unknown. PieA exhibits smeared bands probably due to its basic property (theoretical pI= 8.57).
- (d) (His)<sub>10</sub> pull-down assay for PieA and SidH(1830-2225). Each protein (200  $\mu$ M) was incubated with the indicated subcomplex (100  $\mu$ M) at room temperature for 30 min and mixed with 70  $\mu$ L of Co<sup>2+</sup> resin. The resin was washed two times with a buffer solution containing 20 mM Tris-HCl (pH 7.5) and 100 mM NaCl, and then two times with the same buffer containing additional 10 mM imidazole. Input proteins (I) and Co<sup>2+</sup> resin-bound proteins (R) were visualized on a denaturing gel. RalF and SetA served as a negative and a positive control, respectively.
- (e) Native PAGE analysis for VpdB(11-485). VpdB(11-485) (6  $\mu$ M) was incubated with DotL(590-783)–DotN–IcmSW–LvgA at 1:1 molar ratio. Newly formed protein bands are indicated by triangles. Lines are drawn to distinguish the new protein band from VpdB(11-485) bands.
- (f) Native PAGE analysis for RalF. RalF (6  $\mu$ M) was incubated with DotL(590-783)–DotN–IcmSW–LvgA at 1:1 molar ratio. No detectable interaction is observable.

Throughout the figures, a representative image from more than three replicate experiments are shown.



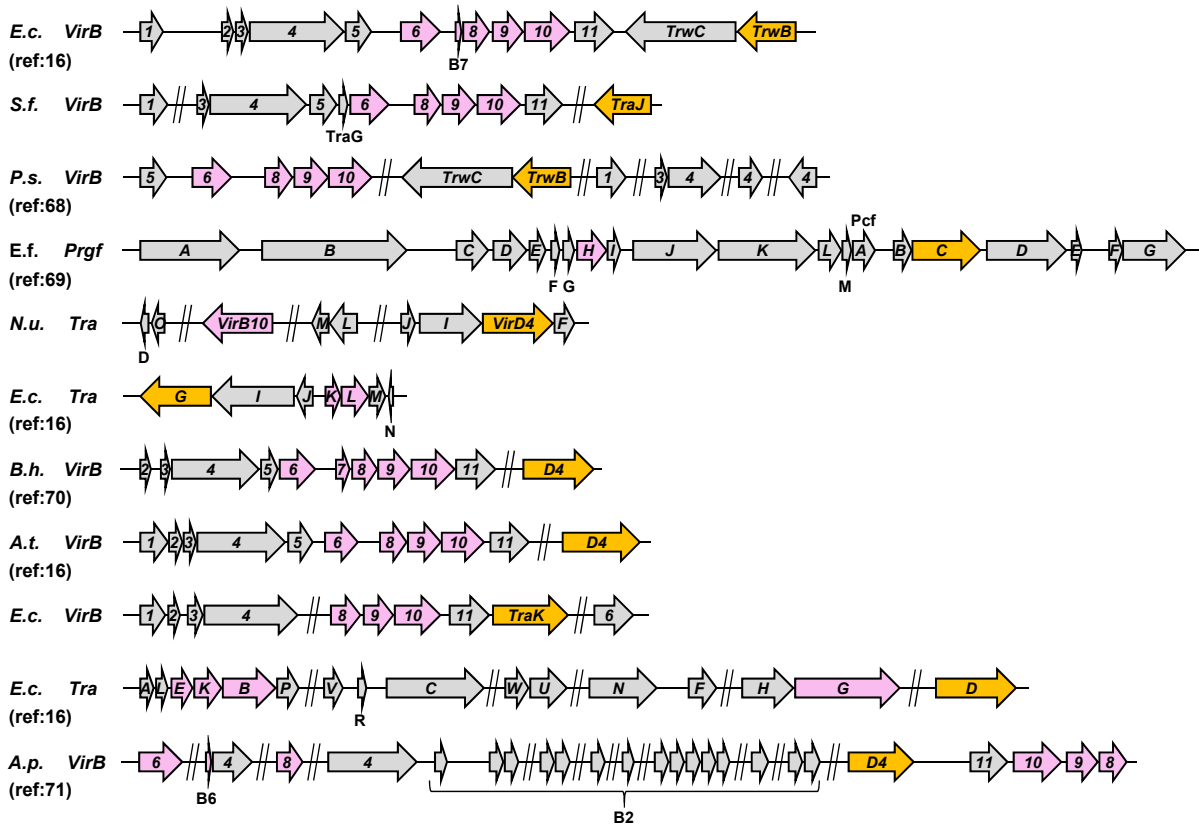


**Supplementary Figure 5. Crystal structure of DotM(161-371)**

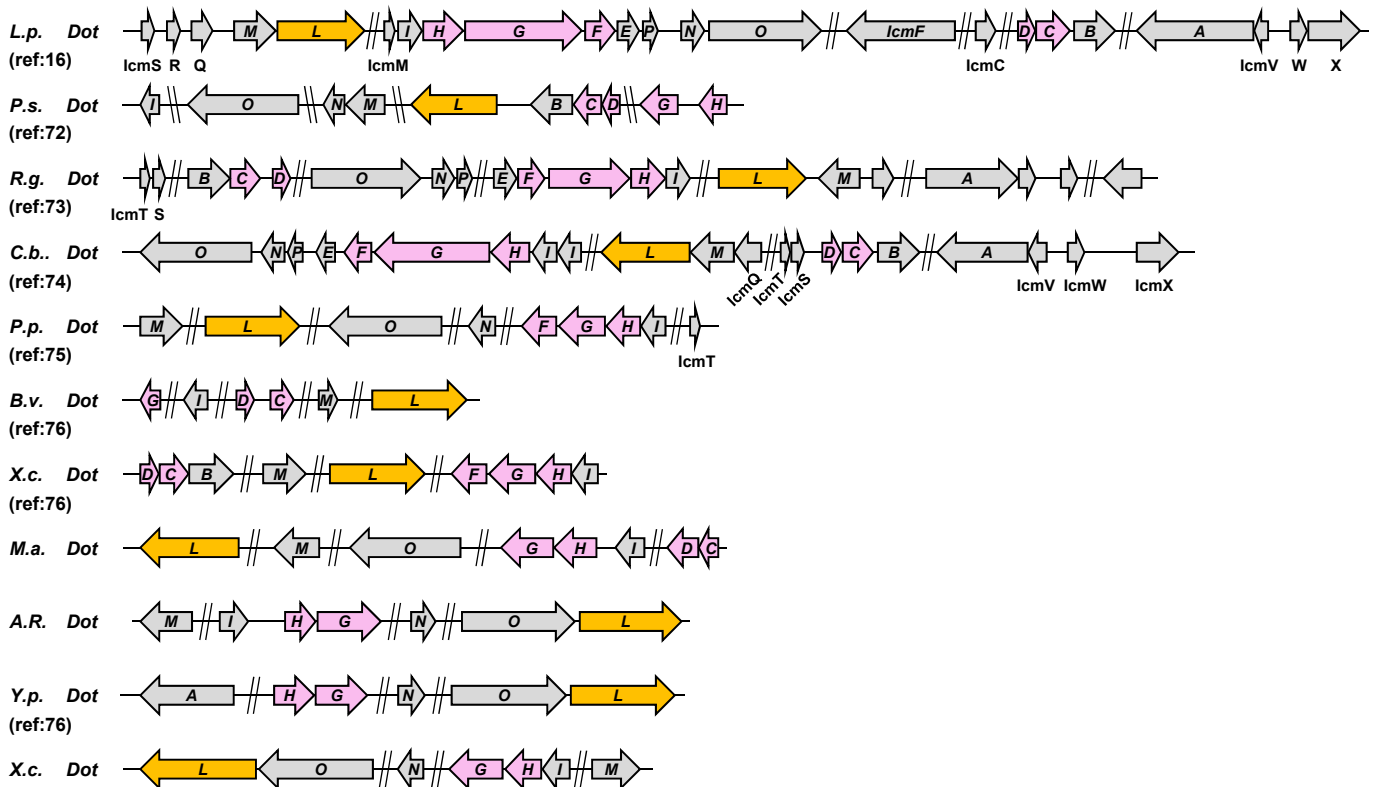
Two perpendicular views are shown. A schematic drawing of the construct is shown at the top. Crystallographic data statistics are summarized in Supplementary Table 1.



## T4ASS genes



## T4BSS genes



## Supplementary Figure 6. Genetic organizations

Shown are the genetic organizations of T4ASS and T4BSS associated with the T4CPs listed in Figure 6. The open-reading frames were identified and annotated by using the RAST server<sup>67</sup>. Homologous genes are color-coded (pink: components of the secretion channel, orange: coupling proteins). The shown references identify between the two subtypes of T4SS. Accession numbers are *E.c.*: *Escherichia coli* (R388 plasmid) (BR000038.1), *S.f.*: *Shigella flexneri 4c* (1205p3 plasmid) (CP012143.1), *P.s.*: *Pseudomonas syringae* (NCPBP880-40 plasmid) (JQ418534.1), *E.f.*: *Enterococcus faecalis* (CF10 plasmid) (AY855841.2), *N.u.*: *Nitrosomonas ureae* (FNUX01000024.1), *E.c.*: *Escherichia coli* (IncP-alpha RP4 plasmid) (X54459.1), *B.h.*: *Bartonella henselae* (JQ701698.1), *A.t.*: *Agrobacterium tumefaciens* (Ti plasmid) (J03320.1), *E.c.*: *Escherichia coli* (O157\_Sal plasmid) (CP001927.1), *E.c.*: *Escherichia coli* (F plasmid) (AP001918.1), *A.p.*: *Anaplasma phagocytophilum* (NZ\_APHI01000002.1), *L.p.*: *Legionella pneumophila* (NZ\_CP013742.1), *P.s.*: *Piscirickettsia salmonis* (CP012413.1), *R.g.*: *Rickettsiella grylli* (NZ\_AAQJ02000001.1), *C.b.*: *Coxiella burnetii* (CP018150.1), *P.p.*: *Pseudomonas putida* (CP003589.1), *B.v.*: *Burkholderia vietnamiensis* (LPCP01000001.1), *X.c.*: *Xanthomonas campestris* pv. *vesicatoria* str. 85-10 (AM039951.1), *M.a.*: *Micavibrio aeruginosavorus* (CP002382.1), *A.* : *Acidovorax* sp. *Root70* (LMHQ01000001.1), *Y.p.*: *Yersinia pseudotuberculosis* (CP000719.1), *X.c.*: *Xanthomonas citri* (CCXZ01000025.1).

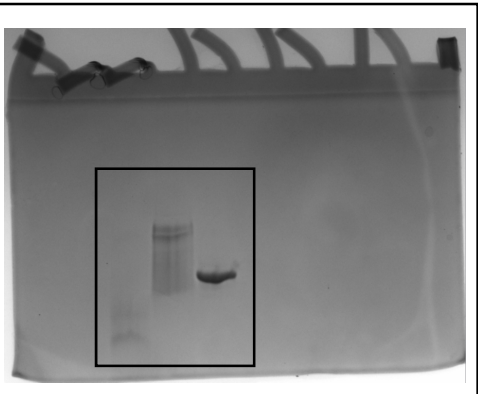
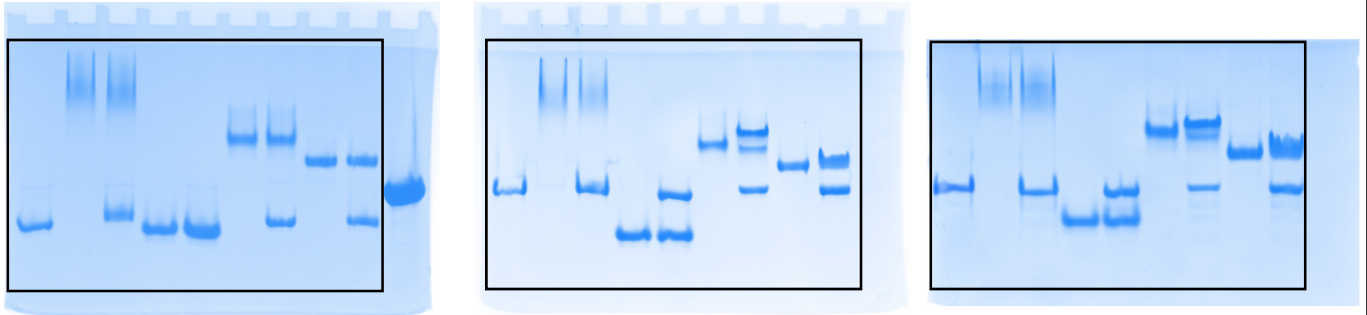


Figure 2a

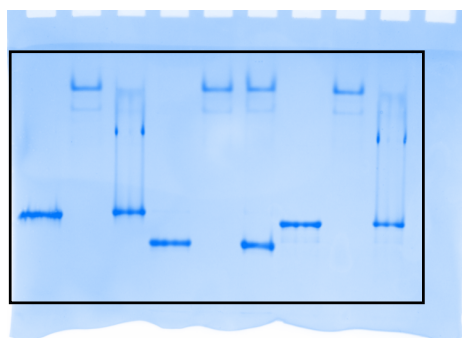
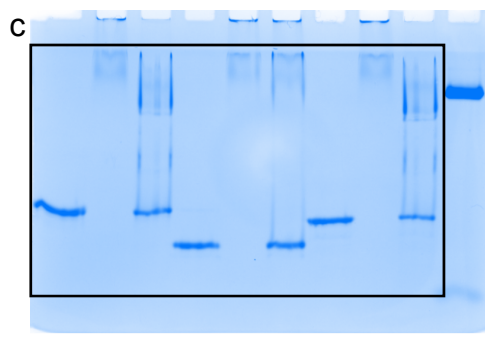
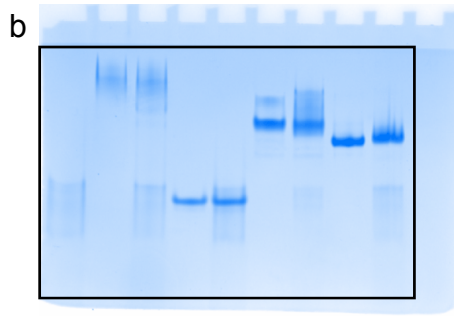
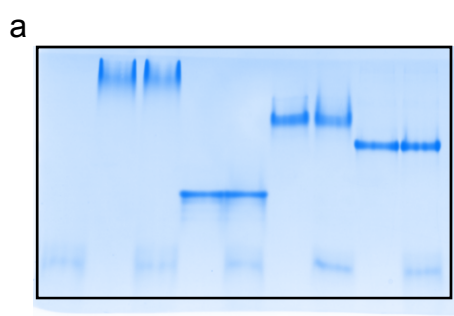


Left

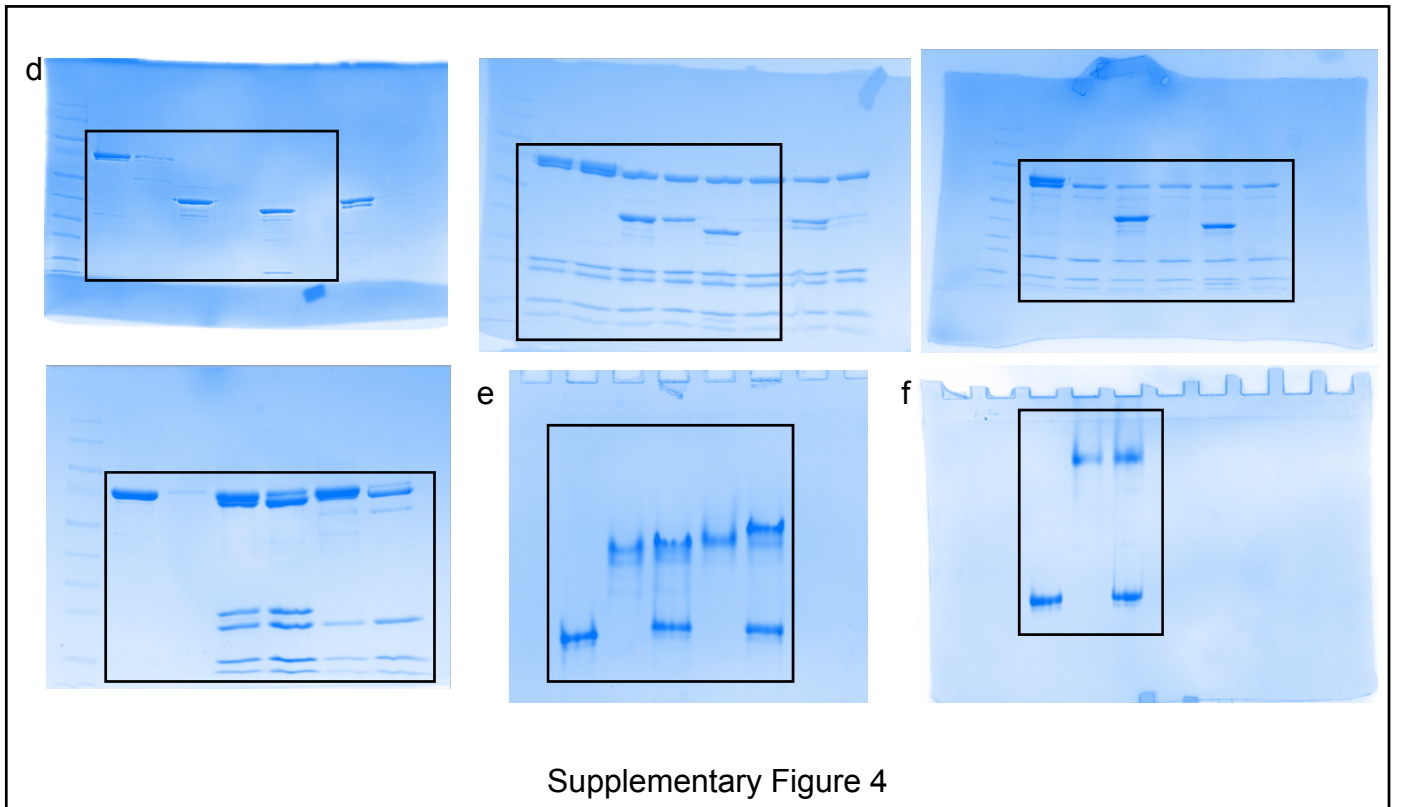
Middle

Right

Figure 4a



Supplementary Figure 4



**Supplementary Figure 7. Full blots of all gels shown in this manuscript**  
Rectangled boxes indicate the regions shown in the indicated figures.

**Supplementary Table 1. X-ray data collection and structure refinement statistics.**

<b>Data Collection</b>	<b>DotL(656-783)–lcmSW (SelMet)</b>	<b>DotN</b>	<b>DotL(590-659)–DotN</b>	<b>DotL(656-783)– lcmSW–LvgA</b>	<b>DotM(161-371) (SelMet)</b>
Space group	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 6 <sub>5</sub> 22	<i>P</i> 2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	<i>P</i> 3 <sub>2</sub>	<i>P</i> 2 <sub>1</sub>
Unit cell dimensions					
a, b, c (Å)	67.597, 75.803, 150.637	155.357, 155.357, 527.711	50.683, 72.220, 170.435	152.325, 152.325, 74.475	50.529, 72.021, 65.691
α, β, γ (°)	90, 90, 90	90, 90, 120	90, 90, 90	90, 90, 120	90, 102.154, 90
Wavelength (Å)	0.9796	1.2828	1.2828	1.0000	0.9796
Resolution (Å)	50-2.0	50-3.0	50-1.8	50-2.8	50-1.8
<i>R</i> <sub>sym</sub>	10(28.9) <sup>a</sup>	10.4(31.7) <sup>a</sup>	6.4(27.2) <sup>a</sup>	8.1(37) <sup>a</sup>	7.9(17.9) <sup>a</sup>
<i>I</i> /σ( <i>I</i> )	28.6(4)	18.8(2.86)	42.2(2.85)	18.8(1.56)	21.3(3.89)
Completeness (%)	92(85.5)	85.4(60.9)	97(91.2)	97.9(90.4)	91(72)
Redundancy	5.4(3.1)	12(1.8)	8.8(3.4)	4.2(2.5)	4.3(2.6)
<b>Refinement</b>					
Resolution (Å)	50-2.0	50-3.0	50-1.8	50-2.8	50-1.8
No. of reflections	86494	109241	104008	46584	70306
<i>R</i> <sub>work</sub> / <i>R</i> <sub>free</sub> (%)	22.1/26.9	25.6/29.7	21.0/25.9	25.0/29.0	17.6/19.8
R.m.s deviations					
bond (Å) / angle (°)	0.008/0.966	0.002/0.408	0.008/0.84	0.01/1.185	0.007/0.83
Average B-values (Å <sup>2</sup> )	16.37	55.68	26.31	64.94	14.69
Ramachandran plot (%)					
Favored / Additional allowed	94.7/5.0	86.3/13.5	92.4/7.6	87.3/12.4	91.5/8.3
Generously allowed	0.2	0.1	0	0.3	0.3
<sup>a</sup> The numbers in parentheses are the statistics from the highest resolution shell.					

## Supplementary References

67. Brettin, T. *et al.* RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Sci Rep* **5**, 8365 (2015).
68. Zhao, Y. F., Ma, Z. H. & Sundin, G. W. Comparative genomic analysis of the pPT23A plasmid family of *Pseudomonas syringae*. *J Bacteriol* **187**, 2113-2126 (2005).
69. Goessweiner-Mohr, N., Arends, K., Keller, W. & Grohmann, E. Conjugative type IV secretion systems in Gram-positive bacteria. *Plasmid* **70**, 289-302 (2013).
70. Schulein, R. & Dehio, C. The VirB/VirD4 type IV secretion system of *Bartonella* is essential for establishing intraerythrocytic infection. *Molecular Microbiology* **46**, 1053-1067 (2002).
71. Rikihisa, Y., Lin, M., Niu, H. & Cheng, Z. Type IV secretion system of *Anaplasma phagocytophilum* and *Ehrlichia chaffeensis*. *Ann N Y Acad Sci* **1166**, 106-111 (2009).
72. Gomez, F. A. *et al.* Evidence of the presence of a functional Dot/Icm type IV-B secretion system in the fish bacterial pathogen *Piscirickettsia salmonis*. *PLoS One* **8**, e54934 (2013).
73. Leclercq, A. & Kleespies, R. G. Type IV secretion system components as phylogenetic markers of entomopathogenic bacteria of the genus *Rickettsiella*. *FEMS Microbiol Lett* **279**, 167-173 (2008).
74. Zamboni, D. S., McGrath, S., Rabinovitch, M. & Roy, C. R. *Coxiella burnetii* express type IV secretion system proteins that function similarly to components of the *Legionella pneumophila* Dot/Icm system. *Mol Microbiol* **49**, 965-976 (2003).
75. Ye, L. M. *et al.* Draft Genome Sequence Analysis of a *Pseudomonas putida* W15Oct28 Strain with Antagonistic Activity to Gram-Positive and *Pseudomonas* sp Pathogens. *Plos One* **9** (2014).
76. Nagai, H. & Kubori, T. Type IVB secretion systems of *Legionella* and other Gram-negative bacteria. *Front Microbiol* **2** (2011).