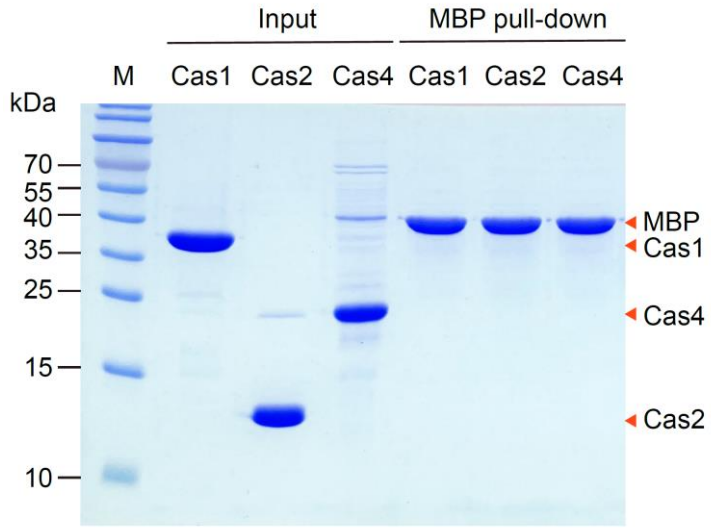

Fig. S1. The MBP only control for MBP pull-down assays. MBP protein does not bind to the tag-free *Synechocystis* Cas1, Cas2, or Cas4 proteins.

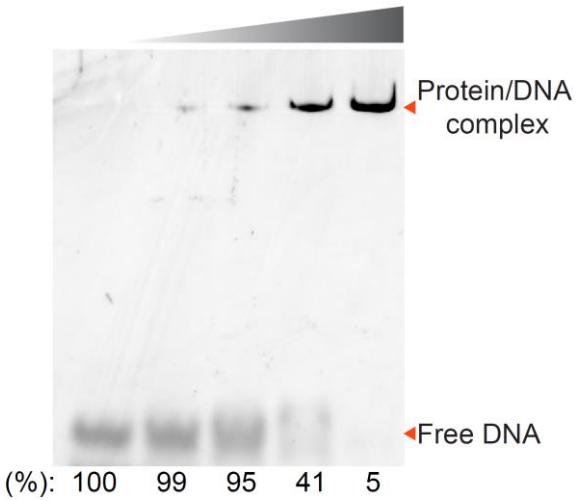

Fig. S2. The binding ability of Ca4-Cas1 complex to the splayed DNA. EMSA is performed using 5' Cy3-labelled splayed prepsapcer DNA (22-bp duplex with 7-nt overhangs). Different concentrations of Cas proteins (0, 0.2, 0.4, 1, and 4 μM) were incubated with 0.2 μM splayed prepsapcer DNA. Percent of the free DNA is calculated based on the gray scanning analysis.

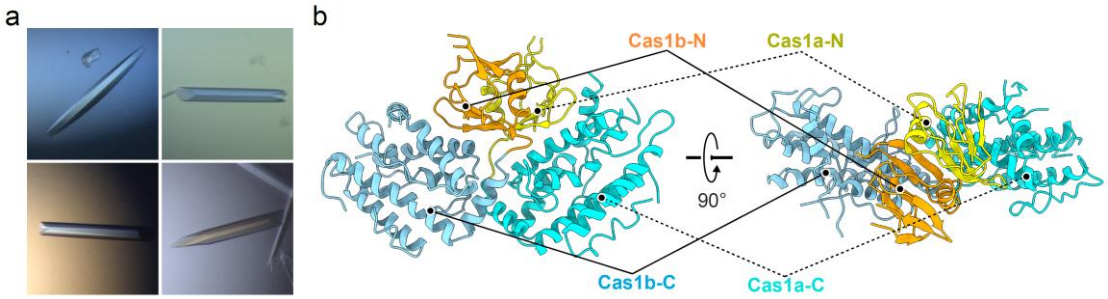

Fig. S3. Crystal Structure of Se-Met Cas1 of *Synechocystis*. (**a**) Crystals of Se-Met Cas1. (**b**) Crystal structure of Se-Met Cas1. Two protomers form a homodimer via N-terminal domains.

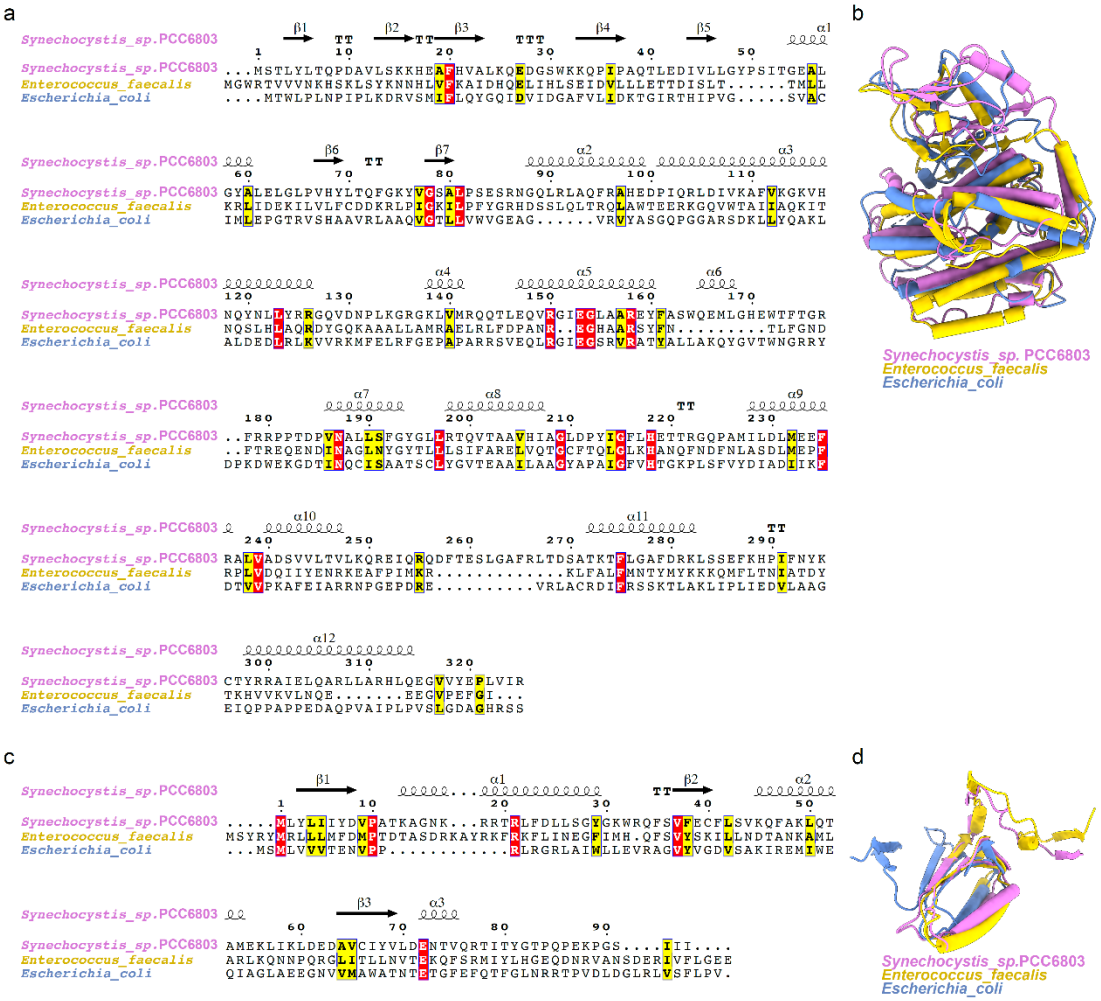

Fig. S4. Alignment of Cas1 and Cas2 from *Synechocystis* sp. PCC6803, *E. faecalis* and *E. coli*. (**a**) Sequence alignment of Cas1 from various species. The secondary structure of *Synechocystis* sp. PCC6803 Cas1 is placed on top. Yellow boxes mark the homologous residues and red boxes mark the identical residues. (**b**) Superposition of Cas1 from various species. (**c**) Sequence alignment of Cas2 from various species. (**d**) Superposition of Cas2 from various species.

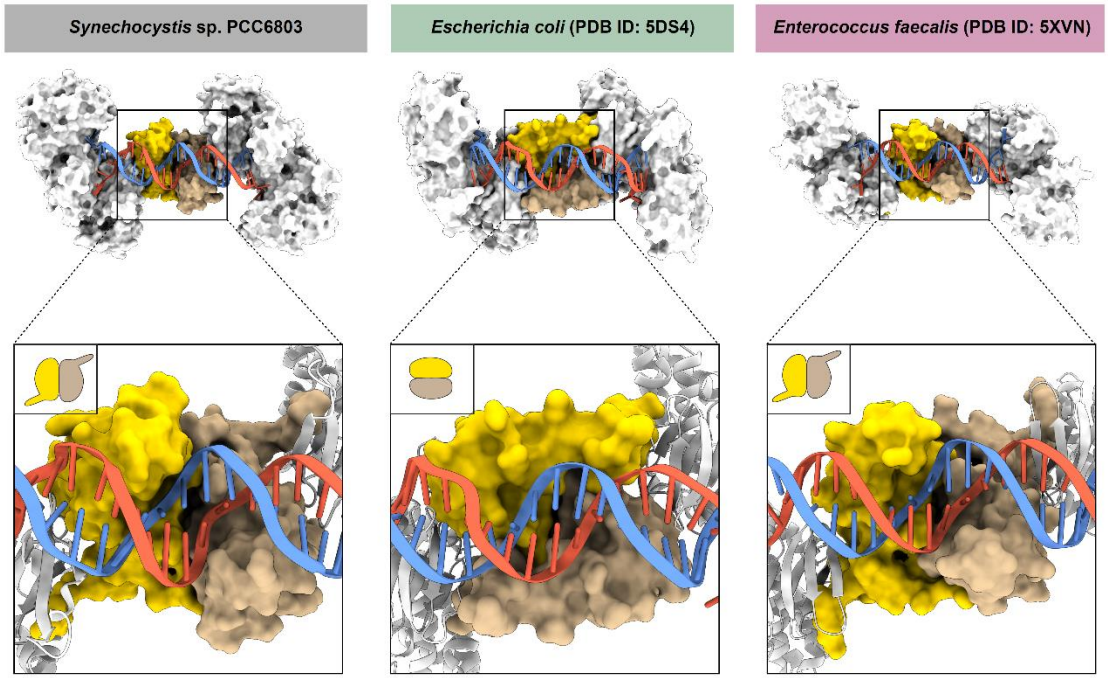

Fig. S5. Comparison of Cas1-Cas2-prespacer Architecture in *Synechocystis* sp.6803, *E. coli* and *E. faecalis*. The Cas1-Cas2-prespacer complexes of *Synechocystis* sp.6803 and *E. faecalis* have a similar Cas2 dimer orientation, which is different from that of *E. coli*.

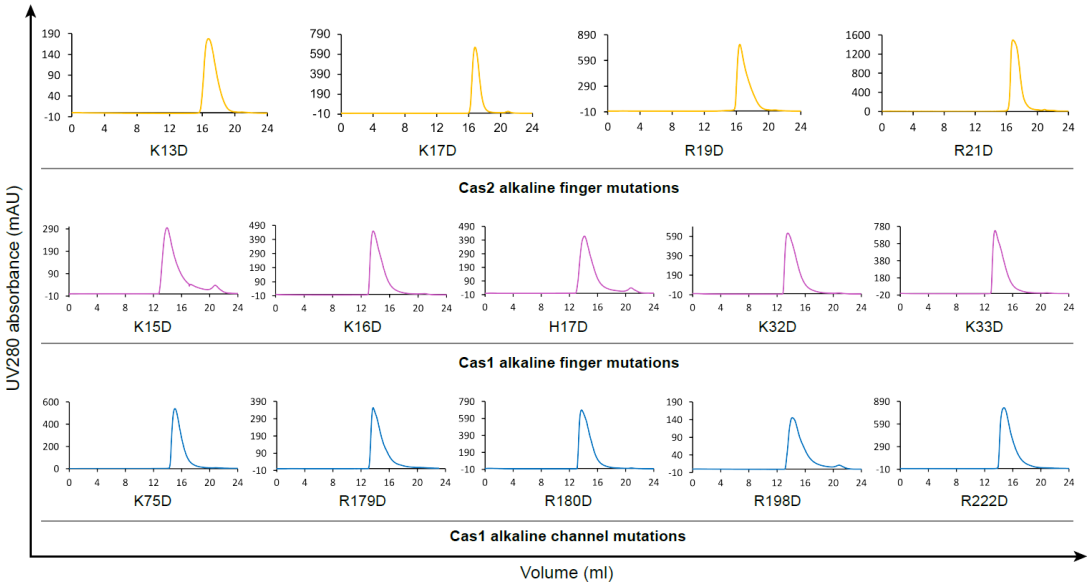

Fig. S6. Gel-filtration Profiles of the Mutants in Lysine Clamp and Arginine Channel. Each mutant is indicated below the gel-filtration profiles.

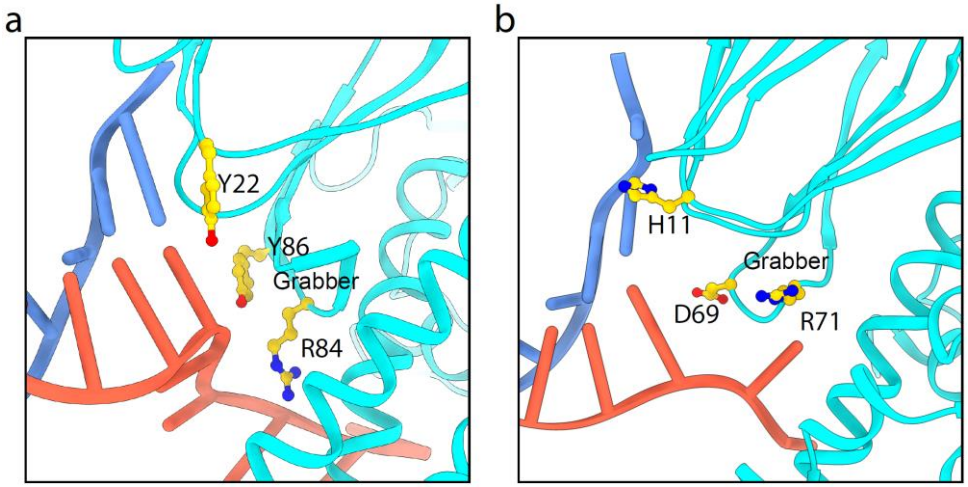

Fig. S7. The Wedge Residues and the Grabber motif in Cas1 of *E. coli* (a, PDB ID: 5DS5) and *E. faecalis* (b, PDB ID: 5XVN). Y22 in *E. coli* and H11 *E. faecalis* function as a wedge to terminate the duplex region of the prespacer. The Grabber motif, Y86 and R84 in *E. coli* and D69 and R71 in *E. faecalis*, bends and guides the single stranded 3'-overhang to the active site of Cas1.

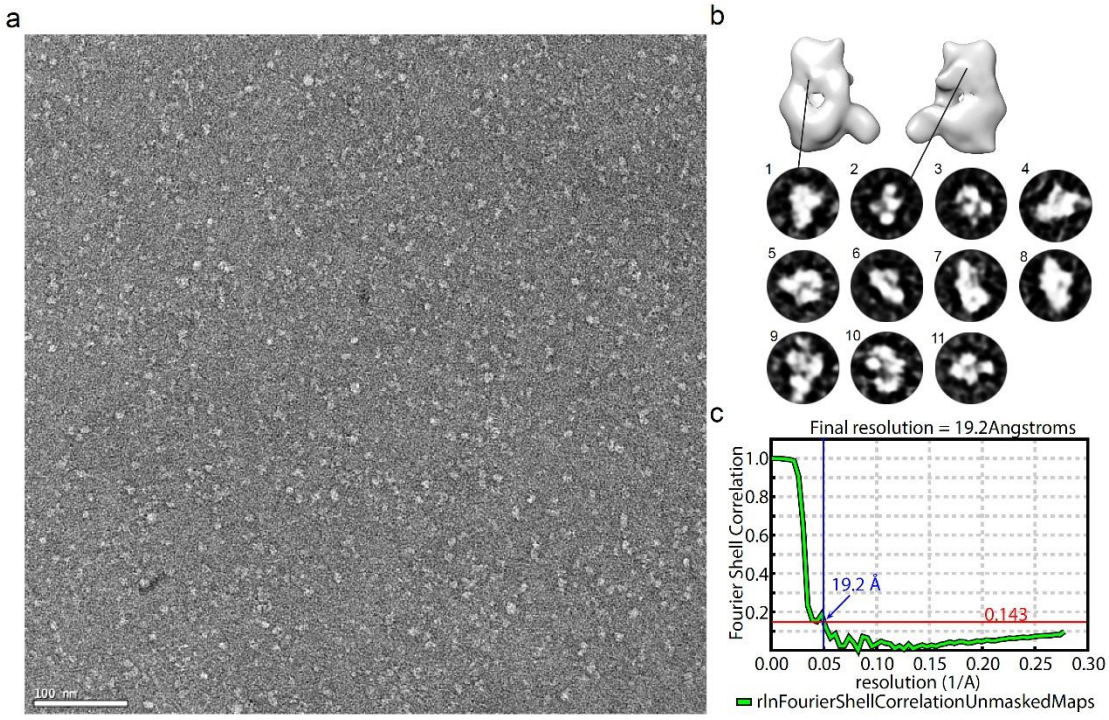

Fig. S8. Structure Reconstruction of the Cas4-Cas1 Complex by Negative Staining Electron Microscopy. (**a**) A representative raw negatively stained micrograph of Cas4-Cas1 complex. (**b**) The two-dimensional class averages of Cas4-Cas1 complex and the surface representation of the 3D reconstruction. (**c**) The Fourier shell correlation (FSC) curve of 3D reconstructed Cas4-Cas1 complex.

Table S1. Data collection and refinement statistics.

	Se-Met Cas1	Cas1-Cas2-22bp	Cas1-Cas2-26bp
Data collection			
Space group	P 6_5 2 2	C 2 2 2_1	P 3_1
Cell dimensions			
a, *b*, *c* (Å)	119.7 119.7 401.7	125. 6 215.2 191.1	185.4 185.4 382.9
α, β, γ (°)	90.0 90.0 120.0	90.0 90.0 90.0	90.0 90.0 120.0
Resolution (Å)	50-2.89(2.95-2.89) *	50-3.70 (3.79-3.70) *	50-3.70 (3.82-3.70) *
R_{pim}	2.8 (49.3)	7.7 (54.2)	10.4 (56.5)
$I/\sigma I$	42.0 (2.0)	11.0 (2.0)	12.7 (2.1)
Completeness (%)	100.0 (100.0)	99.4 (99.4)	99.9(99.9)
Redundancy	73.4 (56.0)	12.5 (9.9)	10.5(10.4)
Refinement			
Resolution (Å)		33.81-3.72	47.56-3.70
No. reflections		25,222	156,966
$R_{work/}$ R_{free} (%)		27.36/32.34	29.57/34.31
No. atoms			
Protein		11,810	71,378
DNA		1,042	7,230
Water		2	0
B-factors			
Protein		40	138
DNA		71	212
R.m.s deviations			
Bond lengths (Å)		0.003	0.004
Bond angles (º)		0.670	0.949

*Highest resolution shell is shown in parenthesis.