Supporting Information

Fidelity of prespacer capture and processing is governed by the PAM mediated interactions of Cas1-2 adaptation complex in CRISPR-Cas type I-E system

Kakimani Nagarajan Yoganand, Manasasri Muralidharan, Siddharth Nimkar, and Baskaran Anand*

Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwahati 781039, INDIA

* To whom correspondence should be addressed.

Tel: +91-361-2582223; Fax: +91-361-2582249; Email: banand@iitg.ac.in

Supporting figures

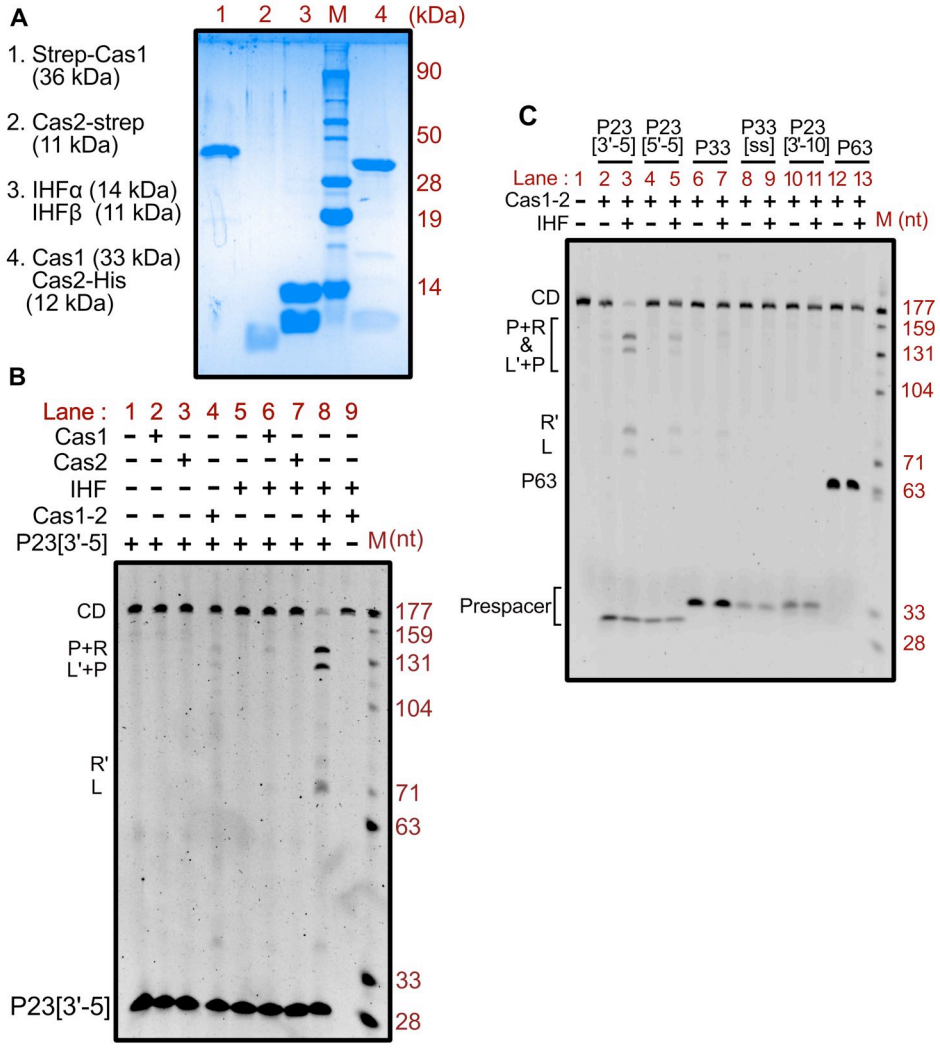

Figure S1. Cas1-2 complex formation necessitates the integration of prespacers into CRISPR DNA. (**A**) 17% SDS-PAGE displays the purified Cas1, Cas2, IHF and Cas1-2. Molecular weight (in kDa) corresponding to proteins in each lane is shown on the left. The molecular weight marker (M) positions are shown on the right. (**B**) Post-stained denaturing gel displaying the samples of spacer integration assay is shown. Absence (–) or presence (+) of Cas1, Cas2, IHF, Cas1-2 and prespacer P23[3'-5] is indicated on top of each lane. Positions of bands corresponding to CRISPR DNA (CD), prespacer (P23[3'-5]) and the DNA fragments that are generated due to prespacer nucleophilic attack and integration (L, R', L'+P and P+R) are displayed. The DNA molecular weight marker (M) positions are shown on the right. Appearance of DNA fragments that are the resultant of prespacer integration in presence of Cas1-2, IHF, prespacer and CD (lane 8) highlights the indispensability of these components. (**C**) Post-stained denaturing gel displaying the samples of spacer integration assay that involves various types of prespacers (Figure 1B) and CRISPR DNA is presented. The type of prespacers utilized in each reaction is depicted on top of the respective lanes. Appearance of DNA fragments that correspond to prespacer integration (L, R', L'+P and P+R) in the samples that contain P23[3'-5] (lane 3), P23[5'-5] (lane 5) and P33 (lane 7) alone, indicates the necessity of prespacer length to be 33 nt.

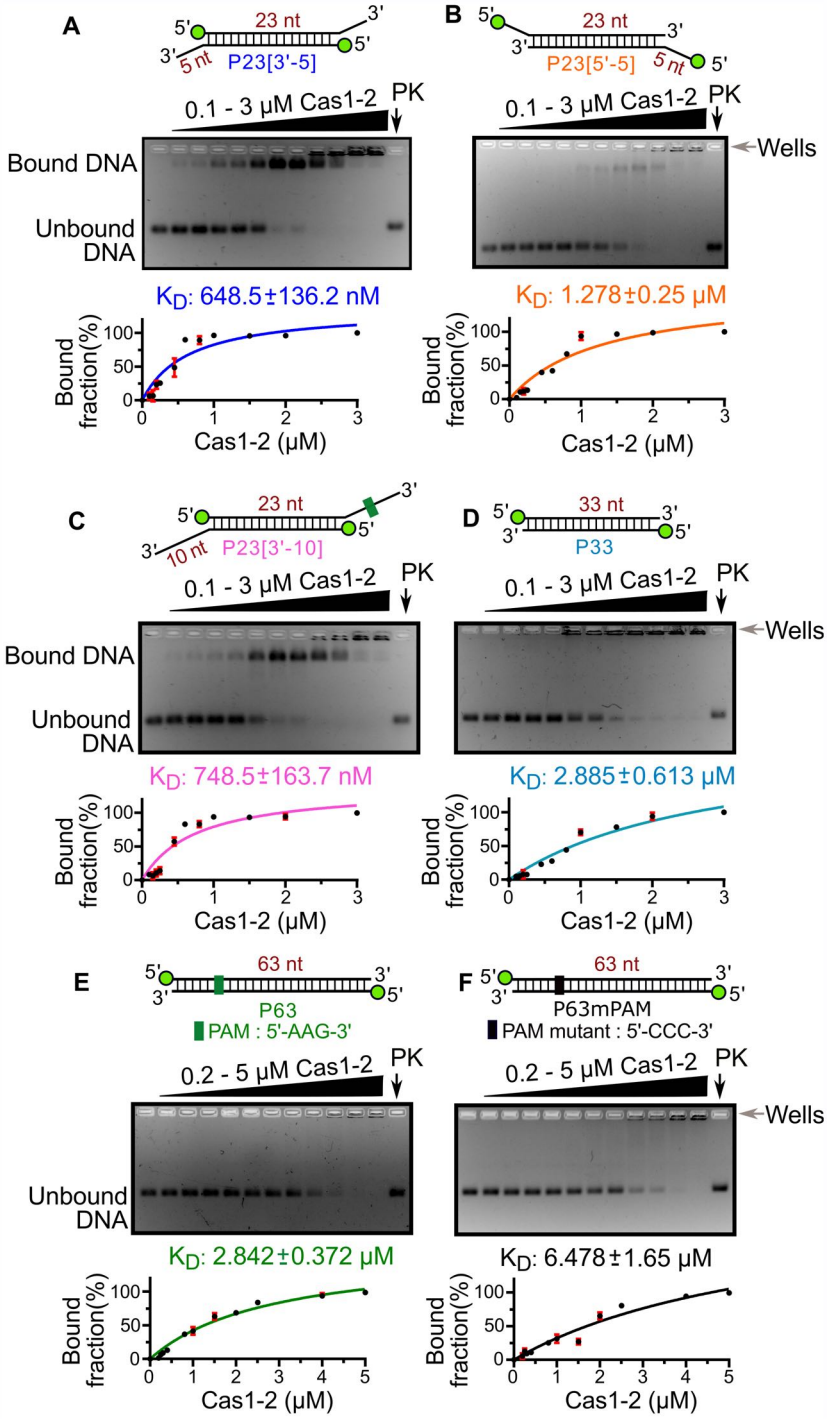

Figure S2. Cas1-2 interacts with prespacers of varied lengths. (**A-F**) Representative agarose gels depicting the interactions of Cas1-2 with 5'-FAM labelled prespacers P23[3'-5] (A), P23[5'-5] (B), P23[3'-10] (C), P33 (D), P63 (E) and P63mPAM (F) are displayed. Schematic representation of the prespacer and the positions of bound and unbound prespacers are shown at the respective gel images. 100 nM of each prespacer DNA (P23[3'-5], P23[5'-5], P23[3'-10] and P33) was incubated with increasing concentrations of Cas1-2 (0, 0.1, 0.15, 0.2, 0.25, 0.45, 0.6, 0.8, 1, 1.5, 2 and 3 μM). In case of prespacers P63 and P63mPAM, 100 nM of DNA was incubated with 0, 0.2, 0.25, 0.3, 0.4, 0.8, 1, 1.5, 2, 2.5, 4 and 5 μM of Cas1-2. The 3'- and 5'- tailed duplex prespacers (A-C) showed a slow migrating Cas1-2-prespacer complex with increasing concentrations of Cas1-2. However, at higher Cas1-2 concentrations, we observed a supershift of DNA in the wells presumably due to the accumulation of DNA-protein aggregates (A-C). In the binding assays that employ blunt prespacers P33, P63 and P63mPAM (D-F), a reduction in prespacer band intensity was observed with increasing concentrations of Cas1-2. In line with previous studies (1), only DNA-protein aggregates were detected in the wells when blunt prespacers were incubated with Cas1-2 (P33, P63 and P63mPAM (D-F)). To further verify the presence of Cas1-2-prespacer complex, aliquots of the samples that contains the mixture of 100 nM prespacer and 3 μM Cas1-2 was treated with proteinase K and the release of intact prespacer was detected (lane PK in A-F). Plots of the bound fraction of prespacer (%) against Cas1-2 concentration (μM) and the estimated K_D±SD values from the binding experiments (in triplicates) are depicted at the bottom of the respective gels in (A-F).

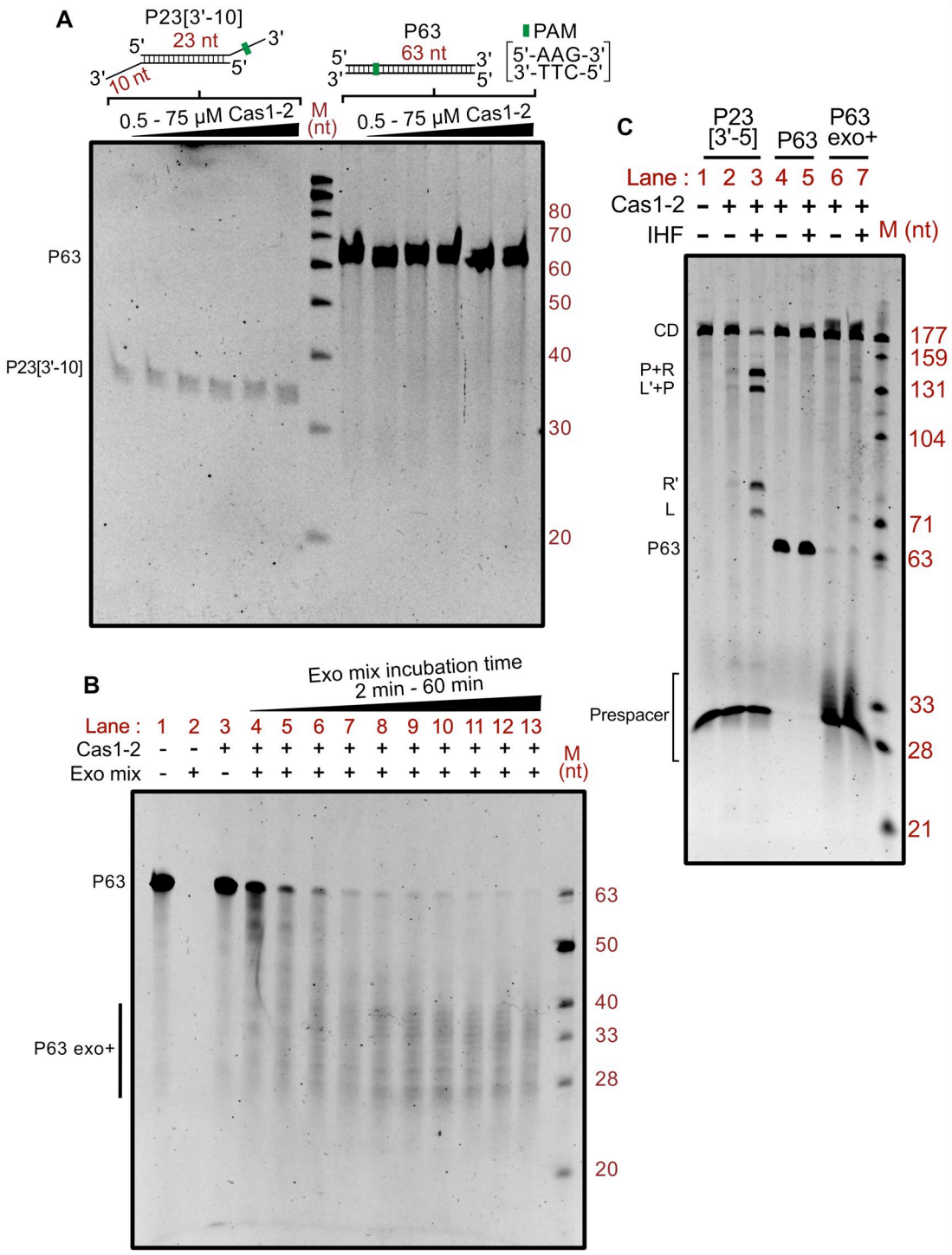

Figure S3. Cas1-2 protects the prespacer boundaries from exonuclease action. (**A**) Denaturing PAGE depicting the interaction of 0.5 µM of prespacer DNA (P23[3'-10] and P63) with increasing concentrations of Cas1-2 (0, 0.5, 5, 10, 25 and 75 µM) is presented. Pictorial depiction of P23[3'-10] and P63 is displayed above their respective lanes. Positions corresponding to each of the prespacer DNA and oligo marker (M) are shown on the sides of the gel. Though P23[3'-10] and P63 were capable of binding to Cas1-2 (Figure S2), generation of smaller DNA fragments was not observed with increasing concentrations of Cas1-2. These observations highlight that Cas1-2 by itself is inept in processing the prespacers. (**B**) Denaturing PAGE depicting the time-dependent nuclease treatment of Cas1-2 bound P63 DNA fragments is displayed. Presence (+) or absence (-) of each reaction component is labelled on top of each lane. Positions corresponding to the substrate (P63) and T5 exo/ExoIII digested DNA fragments (P63exo+) are indicated on the left side, whereas, oligo marker (M) positions are shown on the right. Here the sample containing 0.5 µM of P63 and 6 µM Cas1-2 was treated with exonuclease mixture (T5 exo + ExoIII (3 units each per 20 µl)) for various time points (lanes 4-13: 2,

5, 10, 15, 20, 25, 30, 40, 50 and 60 min). With increase in incubation time, we observed the conversion of P63 to P63exo+ products (~ 33nt). (**C**) Denaturing gel displaying the integration reactions employing various prespacers (P23[3'-5] (lanes 1-3), P63 (lanes 4 and 5) and Cas1-2 protected DNA fragments (P63exo+) (lanes 6 and 7)) is shown. Presence (+) or absence (-) of each reaction component is labelled on top of each lane. Positions of the DNA fragments corresponding to the integration products (L, R', P+R and L'+P) and the prespacers are displayed on the left. DNA molecular weight marker (M) positions are shown on the right. Appearance of DNA bands corresponding to the nucleophilic attack and prespacer ligation upon employing P63exo+ (lane 7) indicates that the DNA regions protected by Cas1-2 during exonuclease digestion could act as prespacer during CRISPR adaptation.

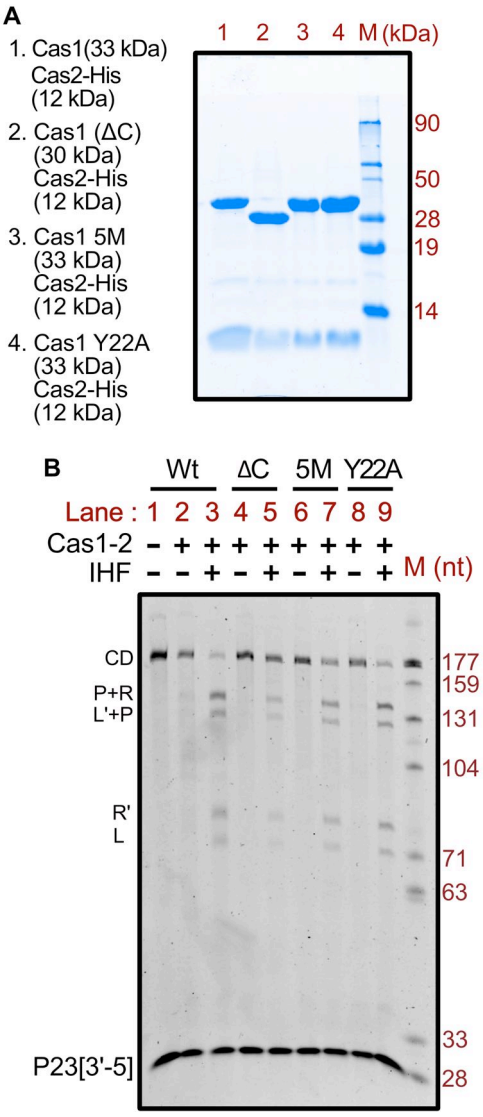

Figure S4. Wt, ΔC, 5M and Y22A display prespacer integration into CRISPR DNA. (**A**) 17% SDS-PAGE displaying purified Wt, ΔC, 5M and Y22A is shown. Molecular weight (kDa) corresponding to proteins in each lane is shown on the left. Protein molecular weight marker (M) positions are shown on the right. (**B**) Ethidium bromide stained denaturing PAGE displaying the samples of spacer integration assay that involves Wt, ΔC, 5M and Y22A variants of Cas1-2 is shown. Absence (–) or presence (+) of Cas1-2 and IHF in integration reaction is indicated on top of the respective lane. DNA positions corresponding to CRISPR DNA (CD), prespacer (P23[3'-5]) and prespacer integration products (L, R', L'+P and P+R) are represented on the left, whereas, DNA molecular weight marker (M) positions are shown on the right side. The appearance of DNA bands corresponding to prespacer nucleophilic attack (L and R') and ligation (L'+P and P+R) in the samples containing IHF and Cas1-2 variant is indicative of integration activity displayed by each variant (Lanes 3, 5, 7 and 9).

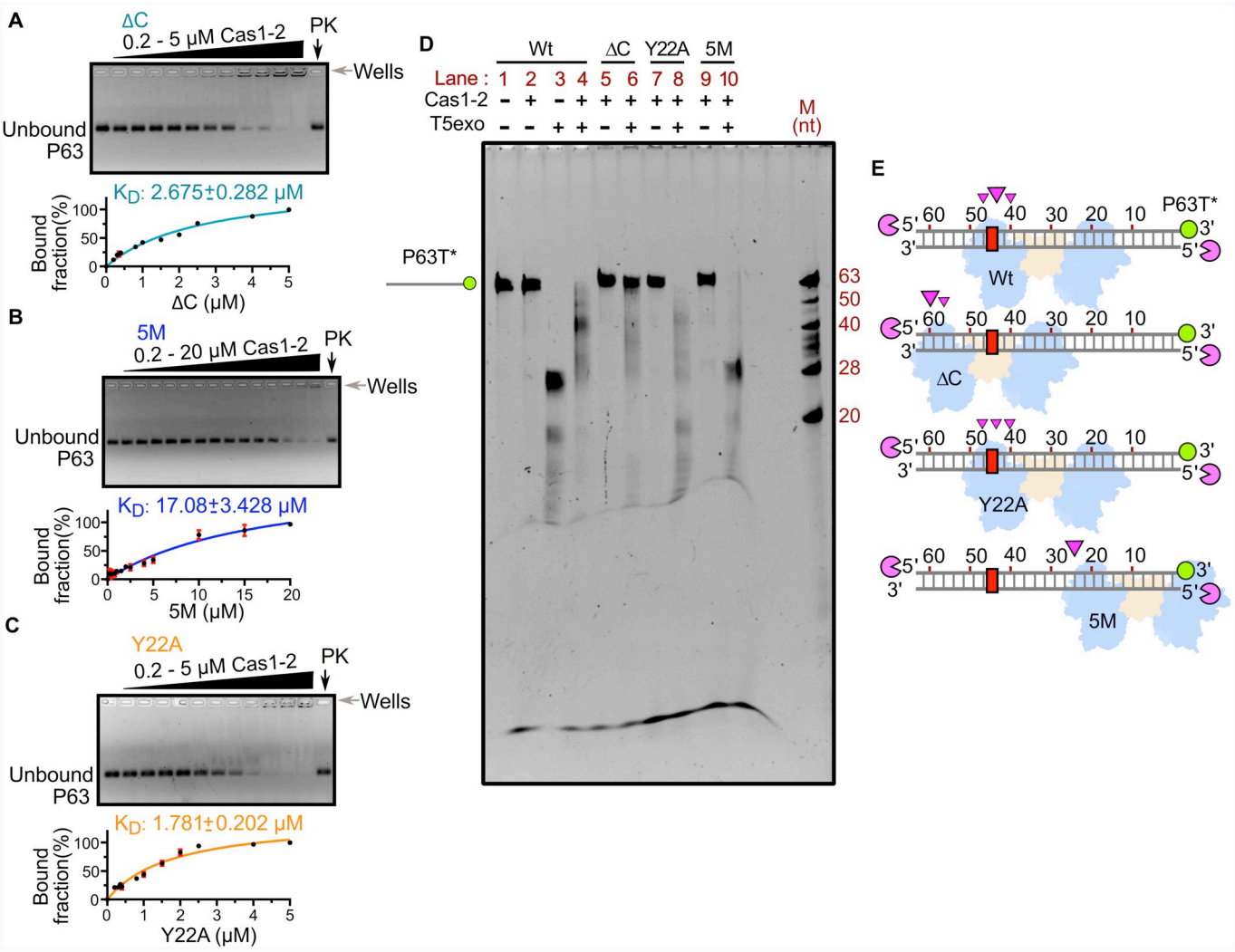

Figure S5. Cas1-2 variants display differing specificities towards prespacers. (**A-C**) Representative agarose gels depicting the interactions of ΔC (A), 5M (B) and Y22A (C) with 5'-FAM labelled prespacer P63 are displayed. The variant of Cas1-2 employed and the position of unbound prespacers are shown in the respective images. 100 nM of P63 was incubated with increasing concentrations of ΔC (A) or Y22A (C) (0.2, 0.3, 0.35, 0.4, 0.8, 1, 1.5, 2, 2.5, 4 and 5 μM) or 5M (B) (0.2, 0.3, 0.35, 0.4, 0.8, 1, 1.5, 2, 2.5, 4, 5, 10, 15 and 20 μM). As observed in Wt-P63 binding assay (Figure S2E), we noted a reduction in P63 DNA intensity and the appearance of DNA-protein aggregates in the wells with an increase in protein concentration. To further verify the presence of Cas1-2-prespacer complex, aliquots of the samples containing a mixture of 100 nM prespacer and 5 μM Cas1-2 variant were treated with proteinase K and the release of intact prespacer was noted (lane PK in A-C). Plots of the bound fraction of prespacer (%) against protein concentration (μM) and the estimated K_D±SD values from the binding experiments (in triplicates) are depicted at the bottom of the respective gels in (A-C). (**D**) Denaturing gel depicting the T5exo treatment of Cas1-2 (Wt (lanes 1-4) or ΔC (lanes 5-6) or Y22A (lanes 7-8) or 5M (lanes 9-10)) bound fluorescein labelled P63T* is displayed. Presence (+) or absence (-) of each reaction component is labelled on top of each lane. Position of P63T* labelled DNA fragment is shown on the left, whereas, oligo marker (M) positions are indicated on the right. (**E**) Schematic illustration depicting the footprinting assay performed in (D) is displayed. DNA substrate P63T* (grey ladder), positions of 3' fluorescein label (green circle) and PAM region (red rectangle) are pictorially represented. Numbering on the DNA represents the distance (in nt) of a particular position from the labelled end. T5 exo (magenta pie) is represented at the susceptible 5'-ends of the DNA substrate. Positions of T5exo stalling points (magenta triangles) and binding sites of each variant of Cas1-2 (Wt or ΔC or Y22A or 5M in blue and brown blobs) that are estimated from nuclease footprinting assay performed in (D) are indicated.

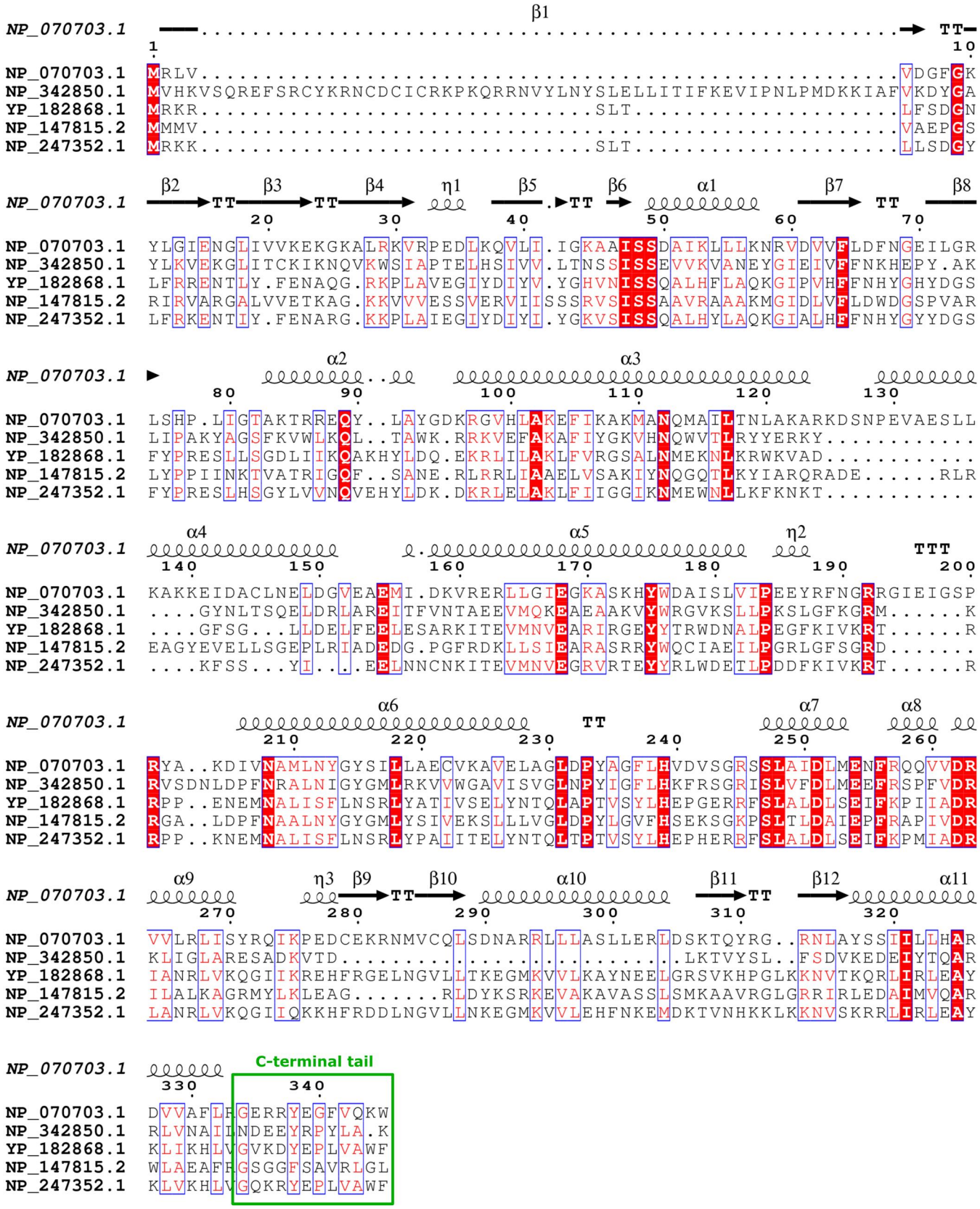

Figure S6. Multiple sequence alignment of Cas1/I-A. A representative sequence alignment of Cas1 derived from *Archaeoglobus fulgidus* DSM 4304 (NP_070703.1), *Sulfolobus solfataricus* P2 (NP_342850.1), *Thermococcus kodakarensis* KOD1 (YP_182868.1), *Aeropyrum pernix* K1 (NP_147815.2) and *Methanocaldococcus jannaschii* DSM 2661 (NP_247352.1) is displayed. A comprehensive sequence alignment of 36 type I-A Cas1 homologs is provided as Supplementary data file F1. Using Cas1 structure (PDB ID: 4N06) from *A. fulgidus* DSM 4304 as a

reference, positions of various secondary structural elements were mapped onto the sequence alignment by ESpript (2). Region corresponding to the C-terminal tail is displayed in green box, whereas, secondary structural features such as alpha helix (α), beta strands (β), 3_{10} helices (η) and beta turns (TT) are depicted at the predicted positions. Amino acid residues that are completely conserved are highlighted in red, whereas, partially conserved residues are boxed and depicted in red font.

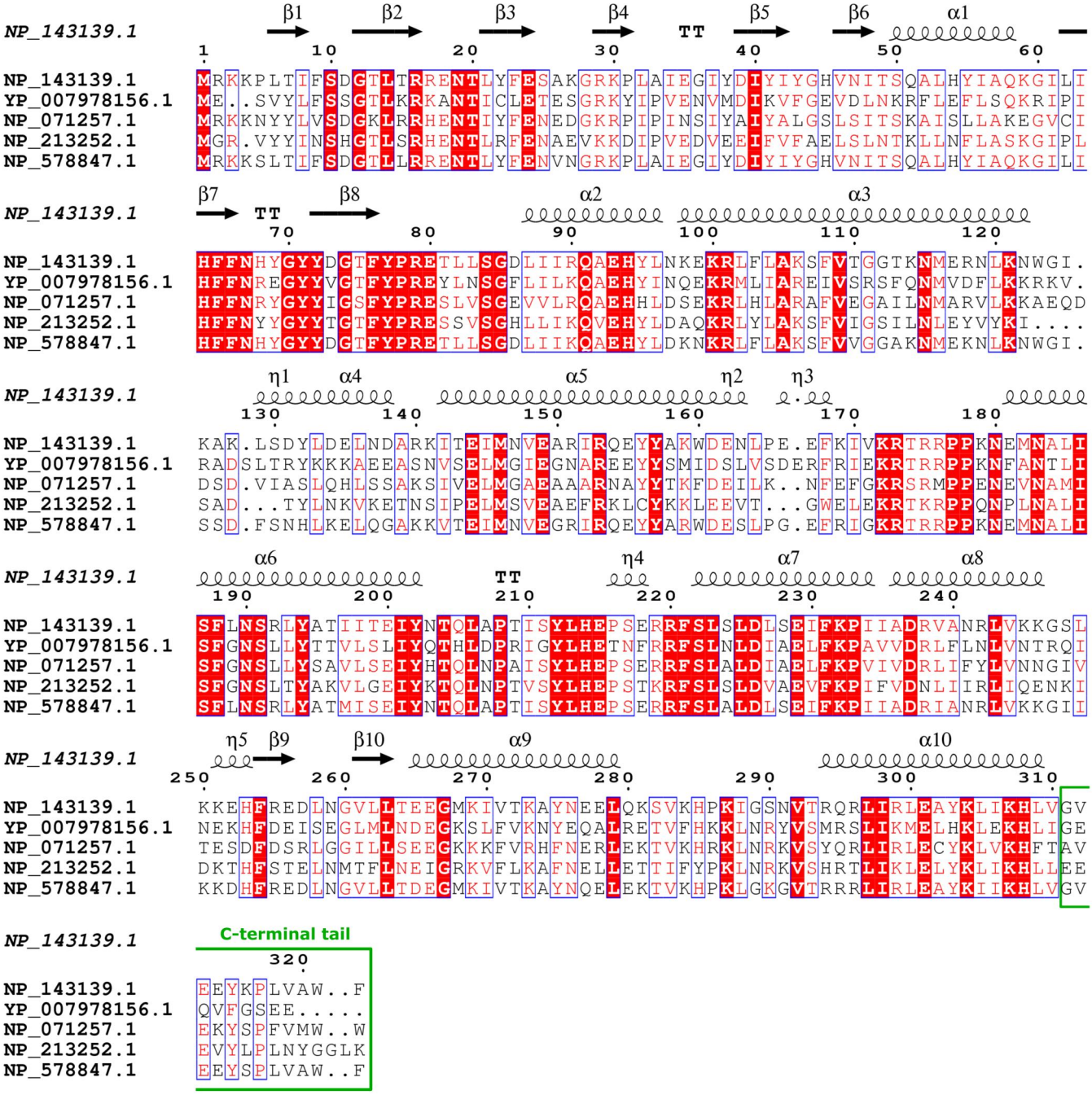

Figure S7. Multiple sequence alignment of Cas1/I-B. A representative sequence alignment of Cas1 derived from *Pyrococcus horikoshii* OT3 (NP_143139.1), *Thermotoga maritima* MSB8 (YP_007978156.1), *Archaeoglobus fulgidus* DSM 4304 (NP_071257.1), *Aquifex aeolicus* VF5 (NP_213252.1) and *Pyrococcus furiosus* DSM 3638 (NP_578847.1) is displayed. A comprehensive sequence alignment of 150 type I-B Cas1 homologs is provided as Supplementary data file F2. Using Cas1 structure (PDB ID: 4WJ0) from *P. horikoshii* OT3 as a reference, positions of various secondary structural elements were mapped onto the sequence alignment by ESpript (2). Region corresponding to the C-terminal tail is displayed in green box, whereas, secondary structural features such as alpha helix (α), beta strands (β), 3_{10} helices (η) and beta turns (TT) are depicted at the predicted positions. Amino acid residues that are completely conserved are highlighted in red, whereas, partially conserved residues are boxed and depicted in red font.

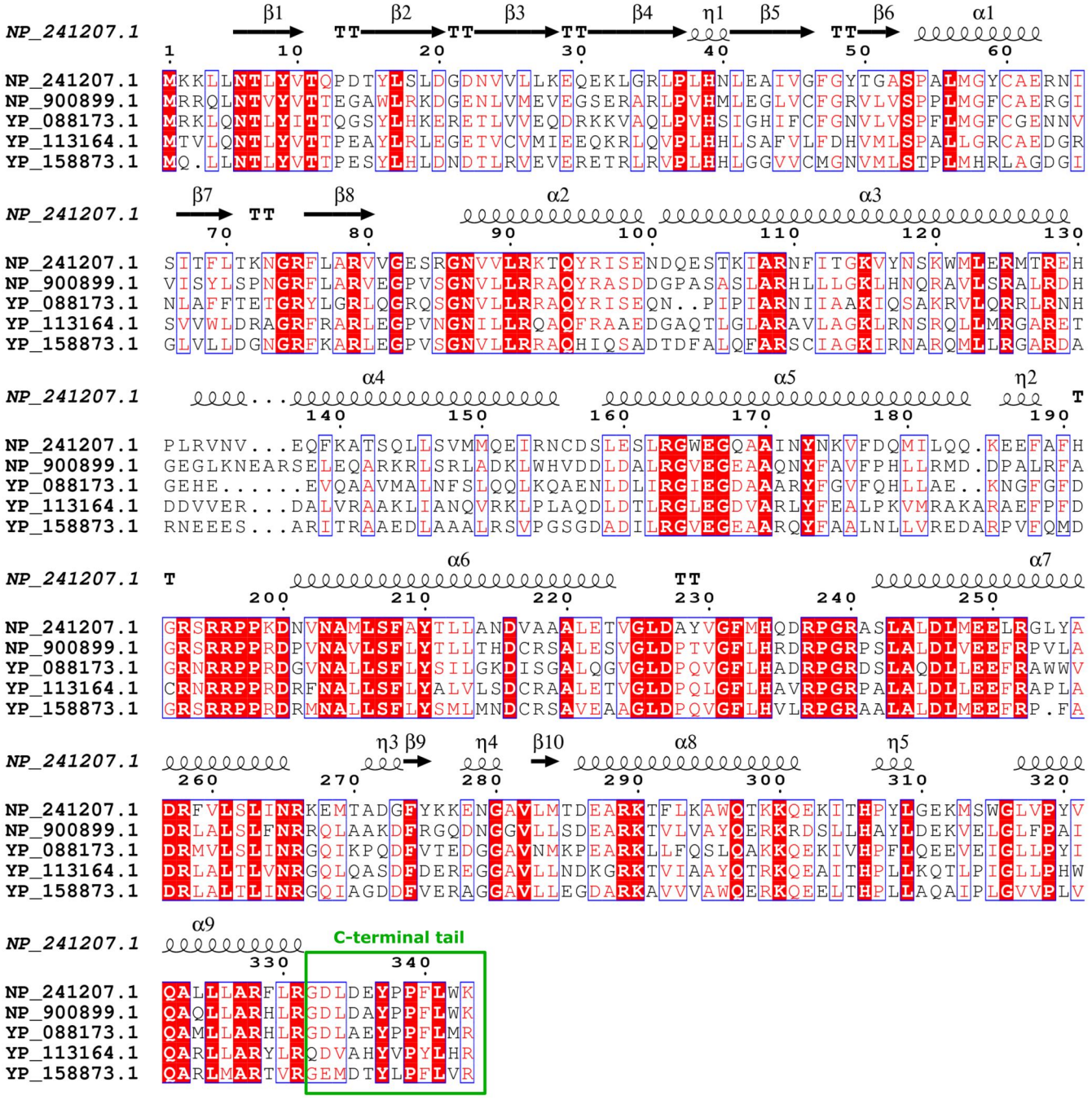

Figure S8. Multiple sequence alignment of Cas1/I-C. A representative sequence alignment of Cas1 derived from *Bacillus halodurans* C-125 (NP_241207.1), *Chromobacterium violaceum* ATCC 12472 (NP_900899.1), *Mannheimia succiniciproducens* MBEL55E (YP_088173.1), *Methylococcus capsulatus* str. Bath (YP_113164.1) and *Aromatoleum aromaticum* EbN1 (YP_158873.1) is displayed. A comprehensive sequence alignment of 129 type I-C Cas1 homologs is provided as Supplementary data file F3. Using Cas1 structure (predicted using I-TASSER) from *B. halodurans* C-125 as a reference, positions of various secondary structural elements were mapped onto the sequence alignment by ESpript (2). Region corresponding to the C-terminal tail is displayed in green box, whereas, secondary structural features such as alpha helix (α), beta strands (β), 3_{10} helices (η) and beta turns (TT) are depicted at the predicted positions. Amino acid residues that are completely conserved are highlighted in red, whereas, partially conserved residues are boxed and depicted in red font.

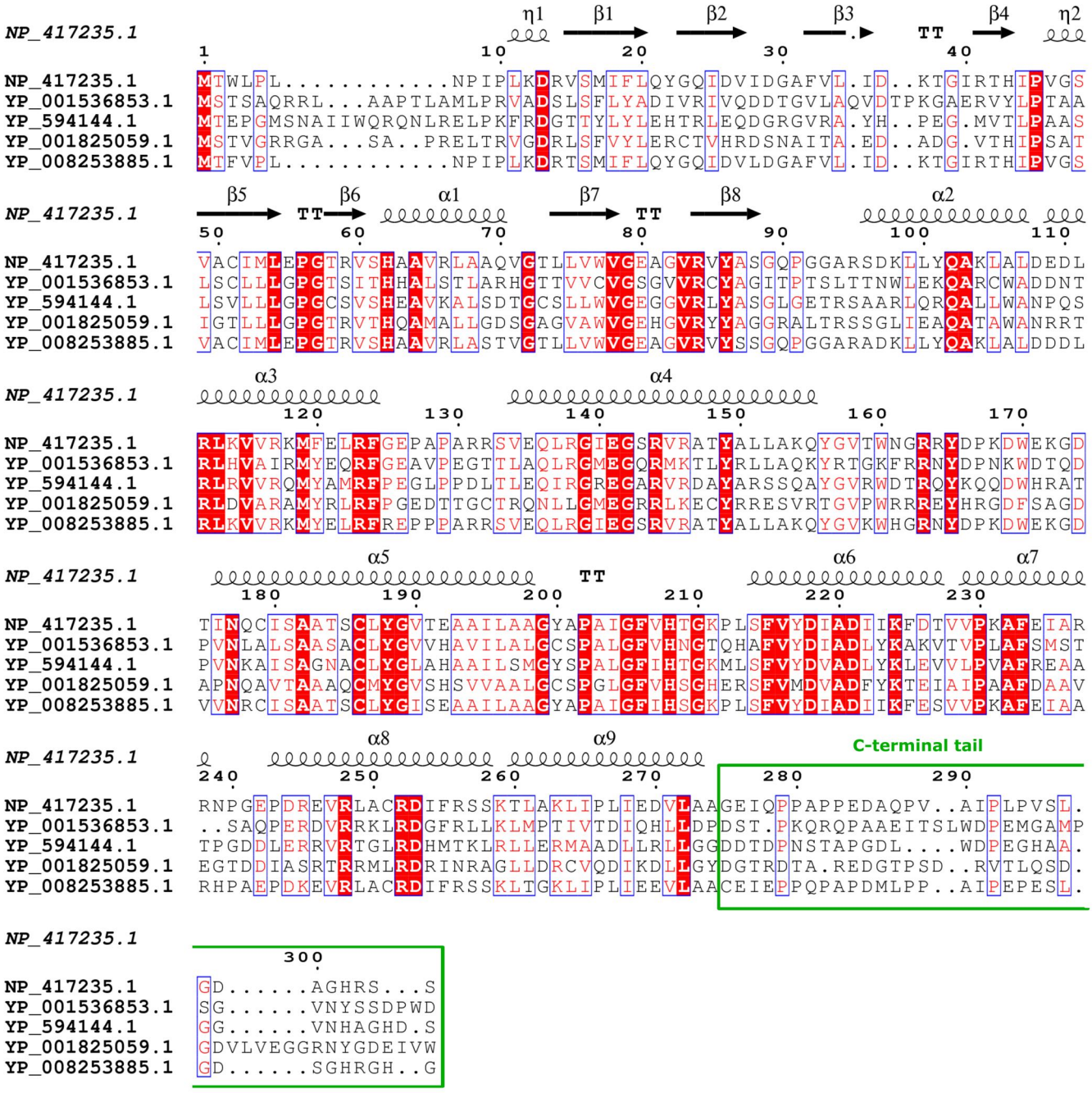

Figure S9. Multiple sequence alignment of Cas1/I-E. A representative sequence alignment of Cas1 derived from *Escherichia coli* str. K-12 substr. MG1655 (NP_417235.1), *Salinispora arenicola* CNS-205 (YP_001536853.1), *Deinococcus geothermalis* DSM 11300 (YP_594144.1), *Streptomyces griseus* subsp. griseus NBRC 13350 (YP_001825059.1) and *Salmonella enterica* subsp. enterica serovar Typhimurium var. 5- str. CFSAN001921 (YP_008253885.1) is displayed. A comprehensive sequence alignment of 116 type I-E Cas1 homologs is provided as Supplementary data file F4. Using Cas1 structure (PDB ID: 5DQZ) from *E. coli* str. K-12 substr. MG1655 as a reference, positions of various secondary structural elements were mapped onto the sequence alignment by ESpript (2). Region corresponding to the C-terminal tail is displayed in green box, whereas, secondary structural features such as alpha helix (α), beta strands (β), 3_{10} helices (η) and beta turns (TT) are depicted at the predicted positions. Amino acid residues that are completely conserved are highlighted in red, whereas, partially conserved residues are boxed and depicted in red font.

Supporting tables

Table S1: Bacterial strains used in the study		
E. coli strain	**Genotype**	**Source**
IYB5101	F⁻Δ(*araD-araB*)*567* Δ*lacZ4787* (::rrnB-3) λ⁻ rph-1 Δ(*rhaD-rhaB*)*568 hsdR514 araB*::T7-RNAp-*tetA*, tet^r	(3)
DH5α	F⁻ Φ80*lacZ*ΔM15 Δ*(lacZYA-argF) U169 recA1 endA1 hsdR17*(r_k^-, m_k^+) *phoA supE44 thi-1 gyrA96 relA1* λ⁻	Invitrogen
TOP10	F⁻ *mcrA* Δ(*mrr-hsdRMS-mcrBC*) ϕ80*lacZ*ΔM15 Δ *lacX*74 *recA1 araD139* Δ(*araleu*)7697 *galU galK rpsL* (str^r) *endA1 nupG*	Invitrogen
BL21(DE3)	F⁻ *fhuA2 [lon] ompT gal (λ DE3) [dcm]* Δ*hsdS λ DE3 = λ sBamHIo* Δ*EcoRI-B int::(lacI::PlacUV5::T7 gene1) i21* Δ*nin5*	NEB

Table S2: Plasmids used in the study		
Plasmid name	**Description**	**Source**
p1R	T7-lac inducible, ColE1 origin plasmid for expressing genes with N-terminally fused Strep-II tag.	Addgene #29664
p13SR	T7-lac inducible, CloDF13 origin plasmid for expressing genes with N-terminally fused Strep-II tag.	Addgene #48328
pMS	T7-lac inducible, RSF origin plasmid for expressing N-terminally 6X His-MBP-SUMO tagged genes.	Addgene #64693
pFGET19_Ulp1	Lac inducible plasmid for expression of 6X His tagged SUMO protease catalytic domain ($Ulp1_{403-621}$).	Addgene #64697
pMut89	CloDF13 origin plasmid for expressing Wt (Cas1-2) under T7-lac inducible promoter and 5M (Cas1(Q24H, P202Q, G241D, E276D, L297Q)-Cas2) under Tetracycline Inducible promoter.	Addgene #80102
pCSIR-T	Plasmid containing CRISPR 2.1 array with 2 repeats	(4)
pCas1-2[K]	Plasmid expressing Cas1 and Cas2 of *E. coli* K-12 MG1655	(4)
p1R-IHFαβ	Plasmid expressing N-terminally Strep-II tagged IHFα and untagged IHFβ under T7-lac promoter.	(5)
p13SR-Cas1	Plasmid expressing N-terminally Strep-II tagged Cas1 under T7-lac promoter.	(5)
pMS-Cas2	Plasmid expressing N-terminally 6X His-MBP-SUMO tagged, C-terminally Strep-II tagged Cas2 under T7-lac promoter.	This study
pCas1-2H	Plasmid expressing Cas1 (Wt), C-terminally 6X His tagged Cas2 under T7-lac promoter.	This study
p5M	Plasmid expressing Cas1 (5M: Q24H, P202Q, G241D, E276D, L297Q), C-terminally 6X His tagged Cas2 under T7-lac promoter.	This study
pY22A	Plasmid expressing Cas1 (Y22A), C-terminally 6X His tagged Cas2 under T7-lac promoter.	This study
pΔC	Plasmid expressing Cas1 (ΔC: ΔP279-S305), C-terminally 6X His tagged Cas2 under T7-lac promoter.	This study

Table S3: Oligonucleotides used in this study		
Oligo name	**Sequence in 5'-3' orientation**	**Purpose**
Ihfα 1R F	TACTTCCAATCCAATGCAATGGCGCTTACAAAAGCTGAAATGT	Generation of cassette encoding IHF with flanking sequences of SspI digested p1R (5).
Ihfα-RBS R	TTGGTCATGGTATATCTCCTTCTTAAAGTTAATTACTCGTCTTTGGGCGAAGC	
RBS-Ihfβ F	ATTAACTTTAAGAAGGAGATATACCATGACCAAGTCAGAATTGATAGAAAGACT	
Ihfβ 1R R	TTATCCACTTCCAATGTTATTATTAACCGTAAATATTGGCGCGATCGC	
Cas1 13SR F	TACTTCCAATCCAATGCAATGACCTGGCTTCCCCTTAATCC	Generation of DNA encoding Cas1 with flanking sequences of SspI digested p13SR (5).
Cas1 13SR R	TTATCCACTTCCAATGTTATTATCAGCTACTCCGATGGCCTGC	
Cas2 MS F	TCACAGAGAACAGATTGGTGGATCCGGAGGTATGAGTATGTTGGTCGTGGTCACTG	Generation of DNA encoding Cas2 with flanking sequences of BamHI and HindIII digested pMS.
Cas2 strep R	TTATTATTTTTCGAACTGCGGGTGGCTCCAAGCGCTAACAGGTAAAAAAGACACCAACCTTAAAC	
MS strep R	CTTTACCAGACTCGAGTGCGGCCGCAAGCTTTTATTATTTTTCGAACTGCGGGTGGC	
CDF wt Cas1 F	AACTTTAATAAGGAGATATACCATGGCCTGGCTTCCCCTTAATCCC	Generation of DNA constructs encoding Cas1-2 variants (Wt, 5M, ΔC, Y22A) with flanking sequences of NcoI and NotI digested pCas1-2[K].
CDF Cas2 His R	TTATTAGTGATGGTGATGGTGATGAGCGCTAACAGGTAAAAAAGACACCAACCTTAAACC	
CDF His R	TTTCTTTACCAGACTCGAGTGCGGCCGCTTATTAGTGATGGTGATGGTGATGAGC	
ΔC Cas1-2 R	TTCAGGTGGGGCCGGCGGCTATTATTGTATTTCTCCAGCGGCAAGC	
ΔC Cas1-2 F	TAATAGCCGCCGGCCCCACCTGAA	
Cas1 Y22A F	TCGCGTCTCCATGATCTTTCTGCAAGCTGGGCAGATCGAT	
2.1 array F	GGAAATGTTACATTAAGGTTGGTGGGTTG	Monitoring the spacer incorporation into the 2.1 CRISPR array during *in vivo* integration assay.
2.1 array R	CGCTCAGAAATTCCAGACCCGATCC	
Leader F	TGCATCTCGAGGCATGCCTGCAGCGGCCGCCAGTGTGATGGATATCTG	Generation of CRISPR DNA substrates for *in vitro* integration assay.
Repeat2 R	GGATCGGAATTCGAGCTCGGTACCCACCTTTGGCTTCGGCTGC	
P23[3'-5] F	ATTTACTACTCGTTCTGGTGTTTCTCGT	These oligos were annealed to prepare P23[3'-5] prespacer.
P23[3'-5] R	AAACACCAGAACGAGTAGTAAATTGGGC	
P23[3'-10] F	ATTTACTACTCGTTCTGGTGTTTCTCGTCAGGG	These oligos were annealed to prepare P23[3'-10] prespacer.
P23[3'-10] R	AAACACCAGAACGAGTAGTAAATTGGGCTTGAG	

P33 F	GCCCAATTTACTACTCGTTCTGGTGTTTCTCGT	These oligos were annealed to prepare P33 prespacer. Whereas, P33 F oligo alone is used as P33[ss] prespacer.
P33 R	ACGAGAAACACCAGAACGAGTAGTAAATTGGGC	
P23[5'5] F	GCCCAATTTACTACTCGTTCTGGTGTTT	These oligos were annealed to prepare P23[5'-5] prespacer.
P23[5'5] R	ACGAGAAACACCAGAACGAGTAGTAAAT	
P63 F	CTCCGCGCTGTAG**AAG**TCACCATTGTTGTGCACGAC GACATCATTCCGTGGCGTTATCCAGCT	These oligos were annealed to prepare P63 prespacer (6). Residues corresponding to the PAM are depicted in green.
P63 R	AGCTGGATAACGCCACGGAATGATGTCGTCGTGCAC AACAATGGTGA**CTT**CTACAGCGCGGAG	
P63mPAM F	CTCCGCGCTGTAG**CCC**TCACCA**C**TGTTGTGCACGAC GACA**C**CA**G**TCCGTGGCGTTATCCAGCT	These oligos were annealed to prepare P63mPAM prespacer. Mutated residues in the oligo are represented in bold.
P63mPAM R	AGCTGGATAACGCCACGGA**C**TG**G**TGTCGTCGTGCA CAACA**G**TGGTGA**GGG**CTACAGCGCGGAG	
P63 3'FAM-F	CTCCGCGCTGTAG**AAG**TCACCATTGTTGTGCACGAC GACATCATTCCGTGGCGTTATCCAGCT **- FAM(3')**	This oligo was annealed with P63 R to generate P63 T*.
P63 3'FAM-R	AGCTGGATAACGCCACGGAATGATGTCGTCGTGCAC AACAATGGTGA**CTT**CTACAGCGCGGAG **- FAM(3')**	This oligo was annealed with P63 F to generate P63 B*.
P63mPAM 3'FAM F	CTCCGCGCTGTAG**CCC**TCACCA**C**TGTTGTGCACGAC GACA**C**CA**G**TCCGTGGCGTTATCCAGCT **- FAM(3')**	This oligo was annealed with P63mPAM R to generate P63mPAM T*.
P63mPAM 3'FAM R	AGCTGGATAACGCCACGGA**C**TG**G**TGTCGTCGTGCA CAACA**G**TGGTGA**GGG**CTACAGCGCGGAG **- FAM(3')**	This oligo was annealed with P63mPAM F to generate P63mPAM B*.

References

1. Radovcic, M., Killelea, T., Savitskaya, E., Wettstein, L., Bolt, E. L., and Ivancic-Bace, I. (2018) CRISPR-Cas adaptation in Escherichia coli requires RecBCD helicase but not nuclease activity, is independent of homologous recombination, and is antagonized by 5' ssDNA exonucleases. *Nucleic Acids Res.* **46**, 10173-10183
2. Robert, X., and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320-W324
3. Yosef, I., Goren, M. G., and Qimron, U. (2012) Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli. *Nucleic Acids Res.* **40**, 5569-5576
4. Diez-Villasenor, C., Guzman, N. M., Almendros, C., Garcia-Martinez, J., and Mojica, F. J. (2013) CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition specificities among CRISPR-Cas I-E variants of Escherichia coli. *RNA Biol.* **10**, 792-802
5. Yoganand, K. N. R., Sivathanu, R., Nimkar, S., and Anand, B. (2017) Asymmetric positioning of Cas1–2 complex and Integration Host Factor induced DNA bending guide the unidirectional homing of protospacer in CRISPR-Cas type I-E system. *Nucleic Acids Res.* **45**, 367-381
6. Shipman, S. L., Nivala, J., Macklis, J. D., and Church, G. M. (2016) Molecular recordings by directed CRISPR spacer acquisition. *Science* **353**, aaf1175