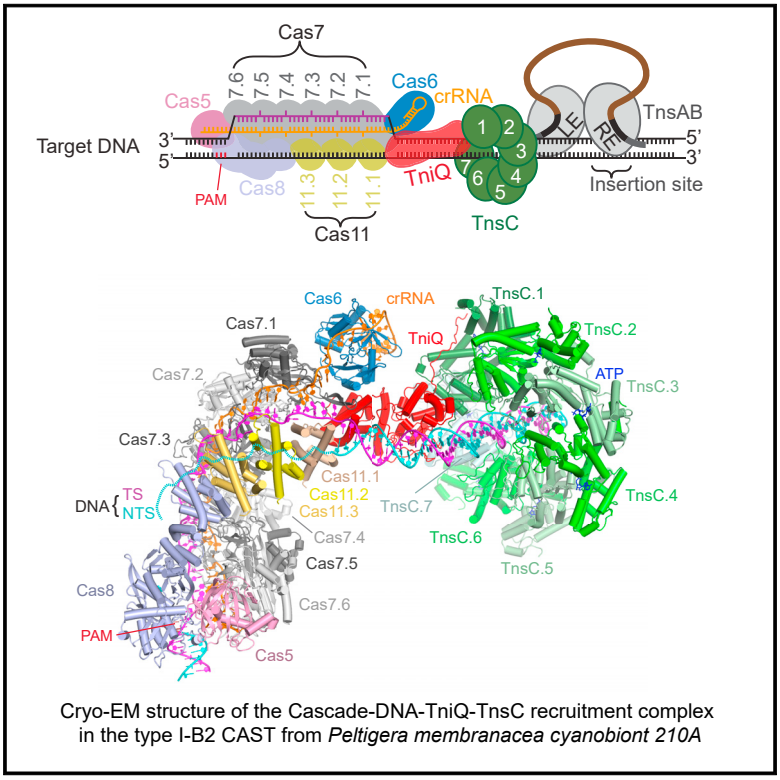
Cell

Molecular mechanism for Tn7-like transposon recruitment by a type I-B CRISPR effector

Graphical abstract



Cryo-EM structure of the Cascade-DNA-TniQ-TnsC recruitment complex in the type I-B2 CAST from *Peltigera membranacea cyanobiont 210A*

Highlights

- Biochemical reconstitution and discovery of Cas11 in type I-B2 *Pmc*CAST
- Cryo-EM structure of the Cascade-DNA-TniQ-TnsC recruitment complex
- Target DNA binding promotes conformational changes in Cascade to recruit TniQ
- TniQ binds to the seam region of the TnsC spiral heptamer

Authors

Shukun Wang, Clinton Gabel, Romana Siddique, Thomas Klose, Leifu Chang

Correspondence

lchang18@purdue.edu

In brief

Cryo-EM and biochemical analyses of a type I-B CAST (CRISPR-associated transposon) provide an improved understanding of various ways in which CAST systems operate; the findings could stimulate ongoing efforts for the development of CASTs as gene knockin tools.

https://www.purdue.edu/newsroom/2025/Q3/blueprints-for-a-molecular-machine-more-powerful-than-crispr/?utm_source=sfmcPT&utm_medium=email&utm_campaign=250909PurdueToday&utm_term=Blueprints+for+a+molecular+machine+more+powerful+than+CRISPR&utm_id=1158912

Wang et al., 2023, Cell *186*, 4204–4215
September 14, 2023
https://doi.org/10.1016/j.cell.2023.07.010

CellPress

CellPress

Cell

Article

Molecular mechanism for Tn7-like transposon recruitment by a type I-B CRISPR effector

Shukun Wang,[1] Clinton Gabel,[1] Romana Siddique,[1] Thomas Klose,[1] and Leifu Chang[1,2,3,*]
[1]Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA
[2]Purdue Institute for Cancer Research, Purdue University, West Lafayette, IN 47907, USA
[3]Lead contact
*Correspondence: lchang18@purdue.edu
https://doi.org/10.1016/j.cell.2023.07.010

SUMMARY

Tn7-like transposons have co-opted CRISPR-Cas systems to facilitate the movement of their own DNA. These CRISPR-associated transposons (CASTs) are promising tools for programmable gene knockin. A key feature of CASTs is their ability to recruit Tn7-like transposons to nuclease-deficient CRISPR effectors. However, how Tn7-like transposons are recruited by diverse CRISPR effectors remains poorly understood. Here, we present the cryo-EM structure of a recruitment complex comprising the Cascade complex, TniQ, TnsC, and the target DNA in the type I-B CAST from *Peltigera membranacea cyanobiont 210A*. Target DNA recognition by Cascade induces conformational changes in Cas6 and primes TniQ recruitment through its C-terminal domain. The N-terminal domain of TniQ is bound to the seam region of the TnsC spiral heptamer. Our findings provide insights into the diverse mechanisms for the recruitment of Tn7-like transposons to CRISPR effectors and will aid in the development of CASTs as gene knockin tools.

INTRODUCTION

CRISPR-associated transposons (CASTs) are mostly Tn7-like transposons that co-opt CRISPR-Cas systems for RNA-guided transposition. Since they were first identified through comparative genomic and phylogenetic analyses,[1] three clades of CASTs have been experimentally characterized, including type V-K,[2] I-F,[3] and I-B systems.[4] As RNA-guided systems, CASTs are promising tools for precise DNA insertion that will have wide applications in biomedical research and the treatment of genetic diseases. In order for CASTs to achieve RNA-guided DNA insertion, CRISPR effectors must recruit multiple transposition proteins. Understanding the molecular mechanisms for the recruitment of these proteins is critical for engineering CASTs into precise DNA insertion tools.

Like prototypic Tn7,[5] CASTs contain defined transposon ends (left end and right end) and encode a heterodimeric transposase (TnsA and TnsB) and a AAA+ ATPase (TnsC). TnsA and TnsB are responsible for cutting the 5′ and 3′ ends of the transposon, respectively, leading to a simple insertion of the transposon into the target site, whereas TnsC regulates the selection of the integration site. When TnsA is absent or inactive, only 3′ ends are cleaved by TnsB and joined to the insertion site, whereas the 5′ ends of the transposon remain linked to the donor. This results in a cointegrate product that contains not only the transposon but also the donor backbone.[6–8] In general, simple insertions are the desired outcome in the context of genome editing.

To disseminate themselves, Tn7-like transposons undergo two modes of transposition: (1) mobilization from the bacterial chromosome to a mobile genetic element that can be horizontally transferred such as a conjugative plasmid and (2), upon entry into a new host, homing from the incoming mobile element into the bacterial chromosome.[4,5] In CASTs, target DNA recognition is achieved by a nuclease-deficient CRISPR effector in an RNA-guided manner. In contrast, target DNA recognition in prototypic Tn7 is accomplished by dedicated proteins (TnsD for homing[9] and TnsE for mobilization[10]). Recent studies suggest that RNA-guided DNA transposition requires TniQ, a homolog of TnsD, and a conserved component in CASTs.[2–4]

Understanding the molecular mechanisms of CASTs not only advances our fundamental knowledge of these systems but also opens up possibilities for the development of CASTs as gene knockin tools. Recently, structural and mechanistic insights into CASTs have been gained through structural studies focusing on the type V-K CAST from cyanobacteria *Scytonema hofmanni* (*Sh*CAST), which utilizes a single Cas12k effector protein for target DNA recognition. A number of structural studies on individual components have been reported, revealing insights into the Cas12k effector,[11,12] TnsC and TniQ,[12,13] and TnsB.[14,15] Furthermore, the structures of a recruitment complex comprising Cas12k-guide RNA, TniQ, TnsC, and the ribosomal protein S15[16] and the holo integration complex[17] have been reported, revealing the host factor S15 and identifying TniQ as a key structural element that bridges Cas12k and TnsC. In parallel, the type I-F CAST from *Vibrio cholerae* (*Vch*CAST) has also received

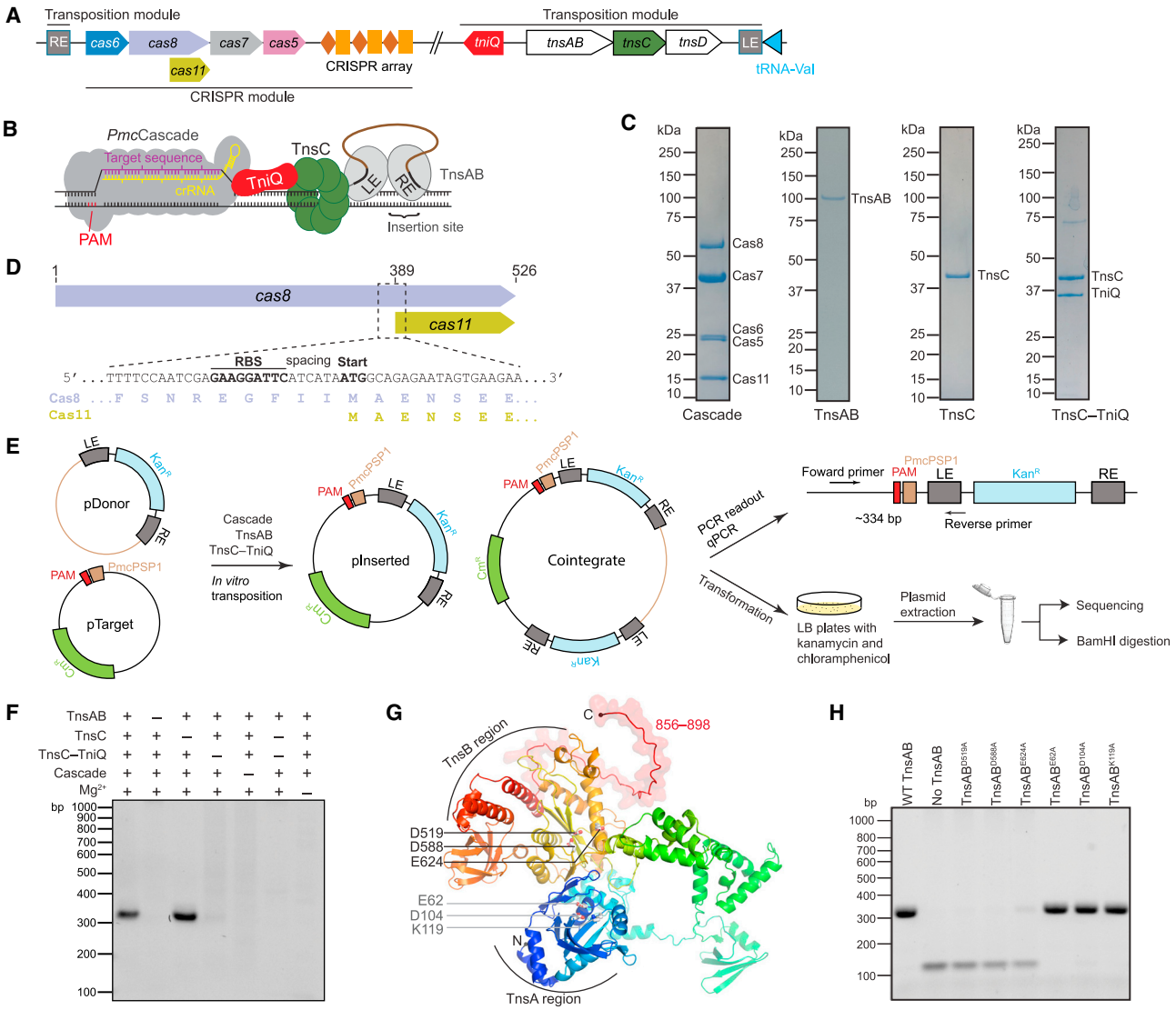


Figure 1. Biochemical reconstitution of *Pmc*CAST

(A) Locus organization of *Pmc*CAST.
(B) Schematic of the components required for RNA-guided transposition in *Pmc*CAST.
(C) SDS-PAGE gels of purified proteins.
(D) Sequence diagram showing the internal translation of Cas11 from the gene of *Cas8*.
(E) Schematic for *in vitro* transposition assay.
(F) PCR readout of *in vitro* transposition assay.
(G) AlphaFold structural prediction of TnsAB with catalytic residues highlighted.
(H) PCR readout of *in vitro* transposition assay using wild-type and mutant TnsAB. See also Figure S1 and Tables S1 and S2.

significant attention. In the type I-F system, a TniQ dimer directly binds to Cascade prior to target DNA binding.[18–21] TnsC in type I-F *Vch*CAST forms a heptamer in the ATP-bound state,[22] although the structural details of the TnsC-TniQ association in this system remain unclear.

To better understand the molecular mechanism of CASTs, it is crucial to investigate additional systems, as this will help elucidate both the similarities and diversity among these systems. The recently characterized type I-B CASTs, including a subtype 1 from *Anabaena variabilis* (*Av*CAST) and subtype 2 from *Peltigera membranacea cyanobiont 210A* (*Pmc*CAST),[4] can perform high-fidelity RNA-guided DNA insertion and are promising candidates for programmable gene knockin.[4] Unlike type V-K and I-F CASTs where both mobilization and homing are RNA guided, type I-B CASTs undergo RNA-guided mobilization and protein-guided homing.[4,23] The type I-B2 CAST *Pmc*CAST is featured by a fused transposase TnsAB (a natural fusion of TnsA and TnsB) (Figure 1A). Mobilization of *Pmc*CAST utilizes the Cascade complex (*Pmc*Cascade hereafter) for RNA-guided target DNA recognition, which then recruits transposition proteins, including TniQ,

TnsC, and TnsAB, to insert the transposon downstream of the sequence targeted by the CRISPR RNA (crRNA) (Figure 1B).

Here, to understand the molecular mechanism by which the *Pmc*CAST Tn7-like transposon is recruited by the type I-B2 CRISPR effector, we used cryoelectron microscopy (cryo-EM) to determine the structures of the *Pmc*Cascade-DNA-TniQ-TnsC recruitment complex and its subcomplexes. We uncovered a mechanism for TniQ recruitment by *Pmc*Cascade upon target DNA binding and how TniQ interacts with the TnsC heptamer. We also revealed detailed interactions between CAST components and the target DNA, ranging from the protospacer-adjacent motif (PAM) sequence to near the insertion site. Our findings highlight the diverse mechanisms of Tn7 recruitment by CRISPR effectors and will facilitate the manipulation of CASTs for genome editing applications.

RESULTS

Biochemical reconstitution of *Pmc*CAST

To biochemically reconstitute the RNA-guided DNA insertion in *Pmc*CAST, we purified *Pmc*Cascade, TnsAB, TnsC, and the TnsC-TniQ complex (Figure 1C). In addition to Cas5, Cas6, Cas7, and Cas8, an unexpected small protein of ~15 kDa was found in purified *Pmc*Cascade samples. Mass spectrometry analysis of the isolated band corresponding to this small protein from the SDS-PAGE revealed peptides within the C-terminal region of Cas8, specifically residues 389–526 (Table S1). Consistently, a sequence analysis of the *cas8* gene in *Pmc*CAST suggested a putative ribosome-binding site (RBS) upstream of the start codon encoding M389 (Figure 1D; Table S2).[24] Internal translation from this site of *cas8* would generate a protein with a theoretical molecular weight of 16.3 kDa, consistent with the unexpected protein. Recent studies showed that Cas11 in some type I systems is expressed from within a large gene, such as *cas10d* in the type I-D system[25] and *cas8* in the type I-C system.[26] Therefore, we assigned this small protein as the Cas11 subunit in *Pmc*Cascade, sharing the same sequence as residues 389–526 of Cas8, most likely resulting from internal translation within the *cas8* gene. Indeed, three copies of Cas11 protein can be modeled into the cryo-EM maps of the *Pmc*Cascade described below. To assess the function of Cas11, we employed the previously established transposition assay in *E. coli*.[4] Mutations to the RBS and start codon of Cas11 in pHelper lead to significantly reduced activity, whereas the deletion of the Cas11-coding region (which inevitably also deletes the C-terminal part of Cas8) results in no detectable RNA-guided transposition activity (Figure S1A). These observations, combined with our structural findings discussed below, which show the involvement of Cas11 in TniQ binding, suggest that Cas11 is likely a crucial subunit of *Pmc*Cascade. TnsAB and TnsC were purified as previously reported[4]; however, the purification of TniQ alone was not successful likely due to inclusion body formation (Figure S1B). To overcome this challenge, we co-expressed TnsC and TniQ and successfully purified the TnsC-TniQ complex (Figure 1C).

Using the purified proteins, a donor plasmid, and a target plasmid, we established an *in vitro* transposition assay (Figure 1E). A typical result is shown in Figure 1F, suggesting successful transposition in an RNA-guided manner when *Pmc*Cascade, TnsC-TniQ, TnsAB, and Mg^{2+} are all present. To examine the exact insertion sites and whether the transposition products are simple insertions or cointegrates, we transformed the products into *E. coli* for amplification. Sanger sequencing of purified plasmids from 18 randomly picked colonies showed that all insertions occurred in an RNA-guided manner, with the insertion sites mostly located between 86 and 93 bp downstream of the PAM sequence (Figures S1C and S1D). Restriction enzyme digestion of the plasmids suggested that 17 colonies (~94%) were simple insertions (Figures S1E and S1F). The ability to perform high-fidelity RNA-guided simple insertion makes *Pmc*CAST a promising candidate for programmable gene knockin.

Simple insertion in *Pmc*CAST is likely dependent on the transposase activity from both TnsA and TnsB, similar to prototypic Tn7.[6] To test this notion, we predicted a structural model of TnsAB in *Pmc*CAST by AlphaFold[27,28] (Figure 1G). Through structural alignments with TnsA in prototypic Tn7 (PDB: 1T0F)[29] and MuA transposase (PDB: 4FCY),[30] which are the two structures with the closest sequence similarity as shown in a HHpred analysis,[31] we identified the residues in TnsAB involved in nuclease activities (Figure S1G). We then made single alanine substitutions to inactivate TnsA (E62A, D104A, and K119A) or TnsB (D519A, D588A, and E624A). *In vitro* transposition assays showed that TnsB activity is essential for RNA-guided transposition, whereas TnsA activity is not (Figures 1H and S1H). However, transposition products are mainly cointegrates when TnsA is mutated (Figure S1I). This is consistent with a previous study on prototypic Tn7, where a single mutation at the active site of TnsA (D114A) switches the transposition mechanism from simple insertion to cointegration.[6] Interestingly, TnsAB contains a C-terminal tail (residues 856–898), the deletion of which almost abolished RNA-guided DNA insertion (Figure S1J).

Cryo-EM analysis of *Pmc*Cascade-DNA-TniQ-TnsC

To understand the molecular mechanism by which the Tn7-like transposon is recruited to *Pmc*Cascade, we assembled the *Pmc*Cascade-DNA-TniQ-TnsC complex for cryo-EM analysis. We first assembled a *Pmc*Cascade-DNA complex using an 85-bp target DNA, including a 10-bp duplex that contains an ATG PAM sequence,[4] followed by a 35-bp spacer that matches the guide RNA and 40 additional base pairs downstream of the spacer. After incubating with excess TnsC-TniQ, the *Pmc*Cascade-DNA-TniQ-TnsC complex was isolated through affinity purification by a polyhistidine-tag in Cas8 (Figure S2A). Our results suggest that the assembly of the *Pmc*Cascade-DNA-TniQ-TnsC complex is dependent on the target DNA, as TniQ cannot be pulled down when the target DNA is omitted (Figures S2B and S2C).

Cryo-EM analysis identified two major populations, accounting for ~45% and ~55% of the screened particles, respectively (Figures S2D–S2F; Table S3). The first population, determined at 2.8 Å resolution, represents a *Pmc*Cascade-DNA complex lacking TnsC and TniQ, where crRNA-DNA hybridization proceeds only to position 23 from the PAM sequence, making 20 crRNA:DNA base pairs (Figures 2A and S2G). We will refer to this complex as *Pmc*Cascade-DNA^{P23} hereafter. The second population, determined at 3.2 Å resolution, represents a

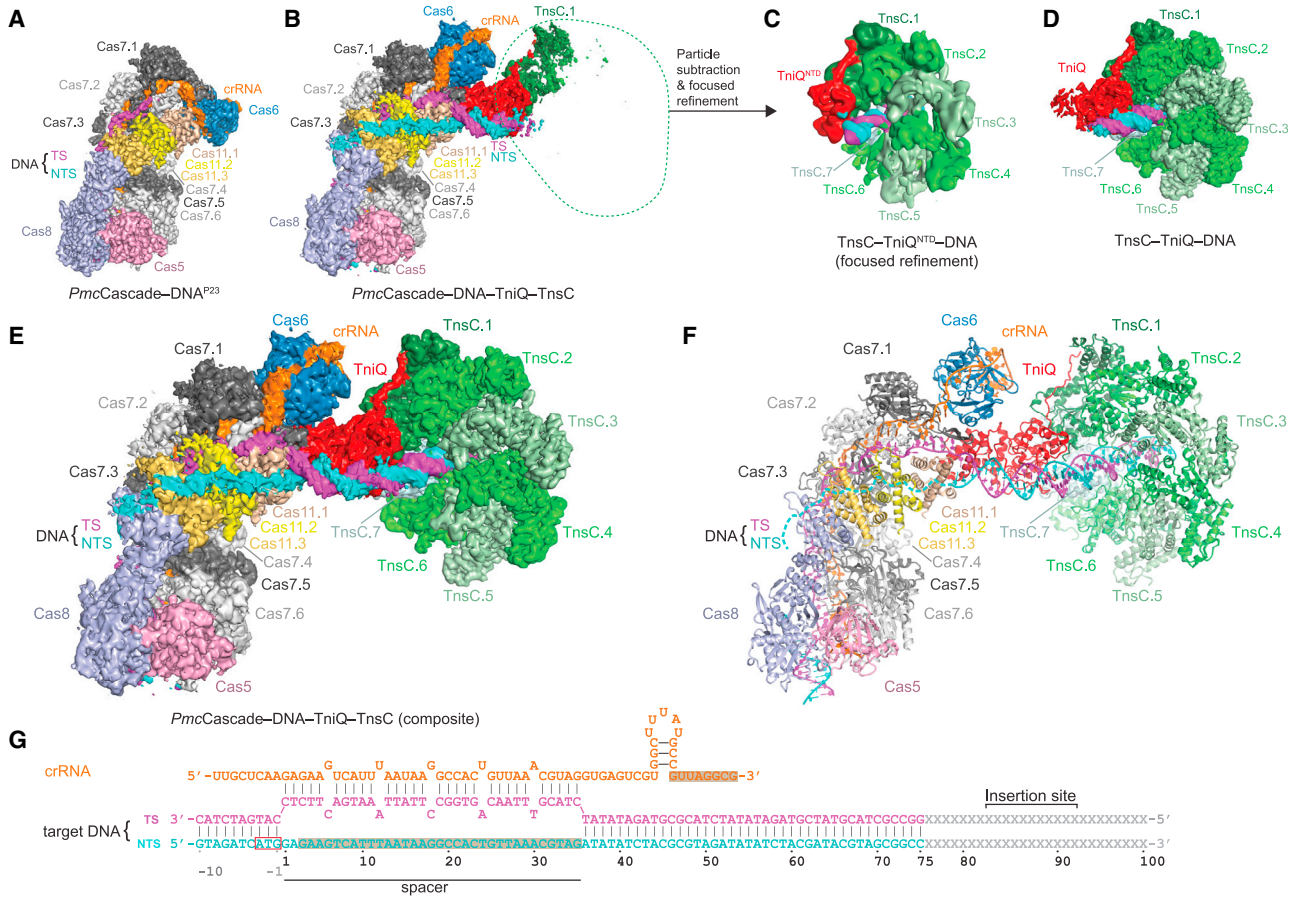


Figure 2. Structure of *Pmc*Cascade-DNA-TniQ-TnsC

(A) Cryo-EM map of *Pmc*Cascade-DNA^{P23} at 2.8 Å.

(B) Cryo-EM map of the *Pmc*Cascade-DNA-TniQ-TnsC complex at 3.2 Å.

(C) Focused refinement map of the TnsC-TniQ region in the *Pmc*Cascade-DNA-TniQ-TnsC complex at ~6 Å.

(D) Cryo-EM map of TnsC-TniQ-DNA at 2.9 Å.

(E) Composite map of *Pmc*Cascade-DNA-TniQ-TnsC by combining (B) and (D).

(F) Atomic model of *Pmc*Cascade-DNA-TniQ-TnsC.

(G) Schematic of target DNA and crRNA. Gray regions are unmodeled in the structure. See also Figures S2 and S4 and Table S3.

*Pmc*Cascade-DNA-TniQ-TnsC complex, where a full R-loop structure is assembled and TnsC-TniQ is bound to *Pmc*Cascade (Figures 2B and S2H). Both maps allowed us to build the atomic models of almost the entire *Pmc*Cascade and the target DNA, except for a few flexible regions such as the non-target strand (NTS) of the target DNA (Figures S3A–S3J). The *Pmc*Cascade-DNA-TniQ-TnsC map further allowed us to build the C-terminal domain of TniQ that binds directly to *Pmc*Cascade via Cas6, Cas7, and Cas11 (Figure S3K). The N-terminal domain (NTD) of TniQ and the TnsC heptamer were determined at an intermediate resolution (~6 Å) (Figures 2C, S2I, and S2J). To obtain a high-resolution structure of these components, we assembled a TnsC-TniQ-DNA complex and determined its structure at 2.9 Å resolution by cryo-EM (Figures 2D, S4A, and S4D–S4F). Taken together, we built a model of the *Pmc*Cascade-DNA-TniQ-TnsC complex, which covers the target DNA from the PAM sequence to a position 75-bp after the PAM, which is ~12 bp away from the insertion site (Figures 2E–2G).

Structure of *Pmc*Cascade-DNA^{P23}

*Pmc*Cascade is assembled by Cas5, Cas6, Cas7, Cas8, Cas11, and the crRNA in a 1:1:6:1:3:1 stoichiometry (Figure 3A). Overall, *Pmc*Cascade as presented in *Pmc*Cascade-DNA^{P23} adopts a double-layered, right-handed spiral structure. The first layer contains six copies of Cas7 (Cas7.1–7.6 from top to bottom) and one Cas5 integrated by the crRNA (Figure 3B). The 35-nt guide region of the crRNA contains six 5-nt segments that base pair with the target DNA and five kink nucleotides between segments (positions 6, 12, 18, 24, and 30), a feature similar to the previously reported type I Cascade.[32–36] The kink nucleotides are stabilized by two tryptophan residues from Cas7 subunits (Figure S3C). The second layer is composed of Cas6 at the top, which is bound to the 3′ hairpin of the crRNA; three Cas11 molecules (Cas11.1–11.3 from top to bottom); and Cas8 (Figure 3B). The C-lobe of Cas8 (residues 389–526; Cas8$^{C\text{-}lobe}$) joins Cas11 to form a Cas11/Cas8$^{C\text{-}lobe}$ filament, which plays an important role in holding the target strand (TS) of the target DNA. The PAM sequence

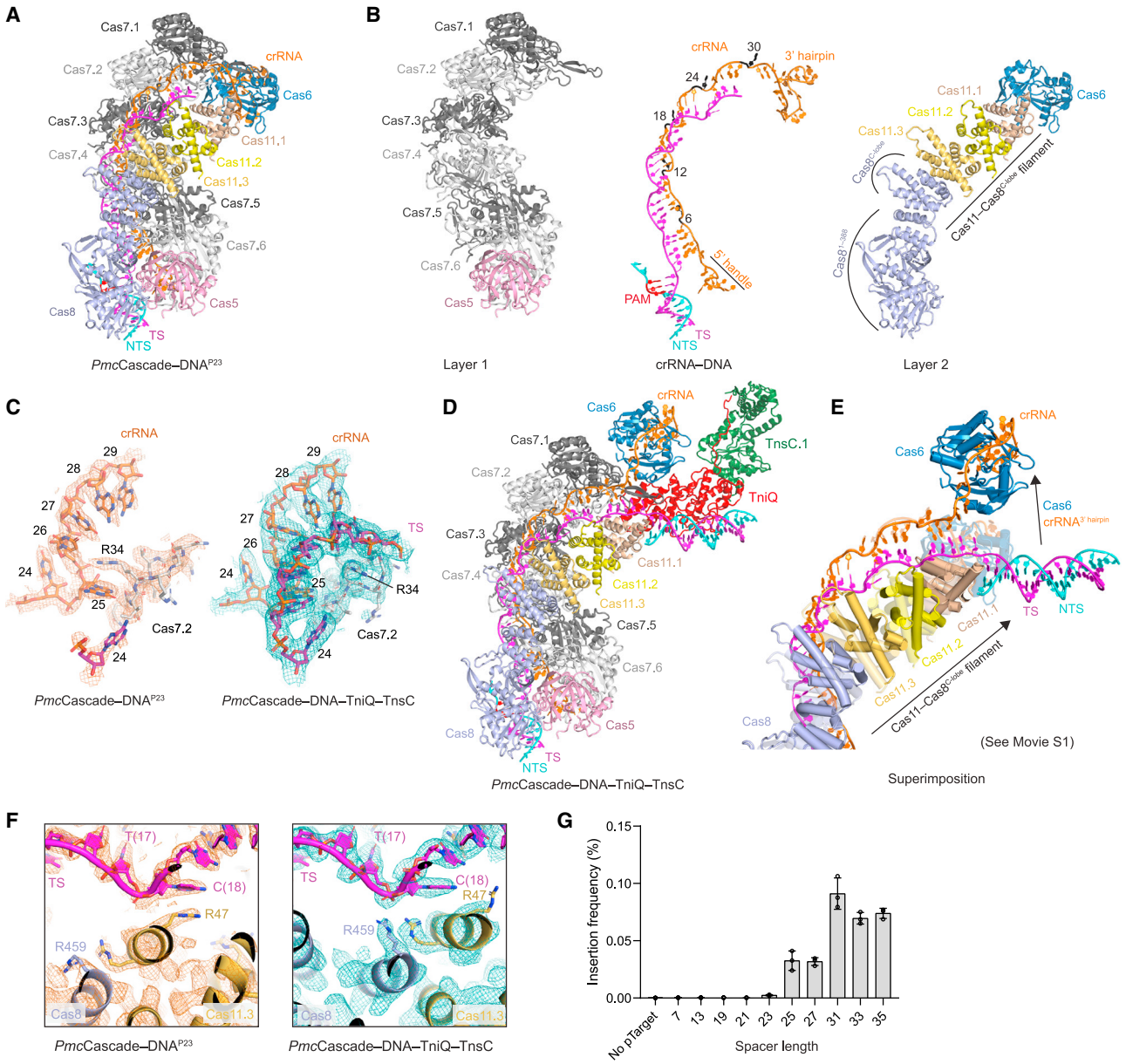


Figure 3. Structure and activation mechanism of *Pmc*Cascade

(A) Atomic model of *Pmc*Cascade-DNA^{P23}.

(B) Atomic model of *Pmc*Cascade-DNA^{P23} shown in three parts: layer 1, crRNA-DNA heteroduplex, and layer 2.

(C) Structures of crRNA-DNA at positions 24–29 in *Pmc*Cascade-DNA^{P23} (left) and *Pmc*Cascade-DNA-TniQ-TnsC (right). Cryo-EM densities are shown in mesh.

(D) Atomic model of *Pmc*Cascade-DNA-TniQ-TnsC.

(E) Conformational changes in *Pmc*Cascade between *Pmc*Cascade-DNA^{P23} (transparent) and *Pmc*Cascade-DNA-TniQ-TnsC (opaque). Cas7 subunits are omitted to highlight the movements in the Cas6-crRNA hairpin and the Cas11/Cas8$^{C\text{-}lobe}$ filament.

(F) Structures of crRNA-DNA at position 18 in *Pmc*Cascade-DNA^{P23} (left) and *Pmc*Cascade-DNA-TniQ-TnsC (right). Cryo-EM densities are shown in mesh.

(G) qPCR analysis for *in vitro* transposition assay using target plasmids with different spacer lengths. Data are presented as mean values ± SD (n = 3). See also Figures S3 and S5 and Video S1.

is located at the interface between Cas8 and Cas5, whereas the crRNA-target DNA heteroduplex is clamped between the two layers (Figure 3B). The region from positions 1 to 23 of the target DNA can be clearly observed in the cryo-EM map, establishing 20-bp with the crRNA (segments 1–5, 7–11, 13–17, and 19–23) (Figure S3B). The map allowed us to trace the target DNA up to position 28, although the cryo-EM densities corresponding to positions 24–28 are relatively weak (Figure S3B). Remarkably, arginine 34 (R34) of Cas7.2 packs against G(25) of the crRNA, disturbing the base pairing between the crRNA and the target

DNA at this position (Figure 3C). A pause of the crRNA-target DNA hybridization at position 23 likely allows *Pmc*Cascade to dissociate or remain inactive when a wrong target is found.

Mechanism for *Pmc*Cascade activation

Comparing *Pmc*Cascade in *Pmc*Cascade-DNA^{P23} and *Pmc*Cascade-DNA-TniQ-TnsC, a major difference is the position of Cas6, which is in a downward conformation in *Pmc*Cascade-DNA^{P23} but in an upward conformation in *Pmc*Cascade-DNA-TniQ-TnsC (Figures 3A and 3D). Cas6 exhibits a ~40 Å movement when comparing the two conformations. Accompanying the profound positional change of Cas6 is an upward movement of the Cas11/Cas8$^{C\text{-}lobe}$ filament by ~8 Å (Figure 3E; Video S1). This movement involves establishing new interactions between the Cas11/Cas8$^{C\text{-}lobe}$ filament and the TS of the target DNA, particularly at the kink positions. For instance, in *Pmc*Cascade-DNA^{P23}, kink C(18) of the TS is packed against R47 of Cas11.3 (Figure 3F). In *Pmc*Cascade-DNA-TniQ-TnsC, the packing interaction is interrupted; concurrently, new electrostatic interaction is established between R459 in Cas8$^{C\text{-}lobe}$ (equivalent to R71 in Cas11) and the phosphate group at the kink. Cas7 and Cas5 in the first layer remain largely unchanged; however, local changes, especially around the newly formed crRNA-target DNA heteroduplex, are observed. For example, R34 in Cas7.2 undergoes a ~180° flip to allow base pairing between the crRNA and the target DNA at position 25 and onward (Figure 3C). Since TniQ-binding sites are not exposed in *Pmc*Cascade-DNA^{P23} (to be discussed below), this structural difference suggests an activation mechanism of *Pmc*Cascade for the recruitment of TnsC-TniQ.

To validate that the downward-to-upward conformation change in Cas6 is induced by the full R-loop formation instead of TniQ association, we determined cryo-EM structures of (1) *Pmc*Cascade alone at 3.6 Å, (2) *Pmc*Cascade bound to the PAM sequence only at 3.1 Å, and (3) *Pmc*Cascade bound to a 65-bp target DNA that contains both PAM and spacer sequences at 3.1 Å (Figures S5A–S5G; Table S3). In both *Pmc*Cascade alone and *Pmc*Cascade-PAM, Cas6 is in a downward conformation (Figures S5H and S5I). In the *Pmc*Cascade-DNA (65 bp) complex, we observed two populations, *Pmc*Cascade-DNA^{P23} and the full R-loop structure (Figures S5D and S5J), with Cas6 in an upward conformation in the full R-loop structure. These structural observations combined with the pull-down assay (Figures S2A and S5B) suggest that the formation of the full R-loop structure triggers the downward-to-upward conformational change in Cas6 and primes TniQ binding.

Motivated by these structural observations, we conducted a mismatch assay to determine the minimum requirement for the length of spacer sequence in RNA-guided transposition. We designed a series of target DNA with increasing matching sequences. *In vitro* DNA transposition assay results indicate that at least a 25-bp target sequence is required for efficient RNA-guided transposition (Figures 3G and S3L). This result is consistent with our structures, which show that crRNA-target DNA hybridization to position 25 is critical in promoting conformational changes in Cas6. We also tested the effects of mismatches at each 5-bp segment. The results suggest that *Pmc*Cascade is more sensitive to mismatches at the PAM-proximal side (Figure S3M), which is consistent with type I Cascade in canonical CRISPR-Cas systems.[37,38]

Structures of TniQ and *Pmc*Cascade-TniQ

TniQ in *Pmc*CAST (*Pmc*TniQ) contains an NTD (residues 1–210) and a C-terminal helical bundle domain (CTD; residues 211–329) (Figures 4A and S6A–S6C). The NTD of *Pmc*TniQ is most similar to TniQ in the *Sh*CAST according to a search in the distance matrix alignment (DALI) server[39] (PDB: 7OXD; *Z* score, 10.6) (Figure S6B). The N-terminal loop (residues 1–20) of *Pmc*TniQ is apparently longer than the N-terminal loop of *Sh*TniQ (residues 1–10),[17] whereas *Vch*TniQ in the type I-F system lacks an equivalent N-terminal loop (Figure S6C). The CTD domain is a specific feature of *Pmc*TniQ, which shares no homology to other structurally known TniQ (Figures S6B and S6D).

The CTD domain of TniQ is bound to *Pmc*Cascade at two sites via Cas6, Cas7.1, and Cas11.1 (Figure 4B). The first binding site is located at the Cas6-Cas7.1 interface through hydrogen bonding and hydrophobic interactions, featured by R245 inserting into a pocket between Cas6 and Cas7.1 and establishing a hydrogen bond with the backbone C=O group (A88) of Cas6 (Figure 4C). At the second binding site, I258, V262, and L265 of TniQ contact Cas11 through hydrophobic interactions (Figure 4D). The R245A mutant of TniQ shows significantly decreased activity compared with wild-type TniQ in the *in vitro* transposition assays (Figures 4E and S6E), whereas we were not able to purify TniQ (in the format of TnsC-TniQ) when I258, V262, and L265 were substituted by glutamic acids. To be noted, the binding sites for TniQ are only available when Cas6 is in the upward conformation. In the Cas6 downward conformation, the first binding site is not formed because the regions in Cas7.1 and Cas6 that create the binding pocket are separated; additionally, the second binding site on Cas11 is covered by Cas6 (Figure 4F).

Structure of TnsC-TniQ-DNA

In the TnsC-TniQ-DNA complex, seven TnsC molecules form a right-handed spiral structure with a seam between the first (TnsC.1; the TnsC molecule that binds to TniQ) and the last (TnsC.7) TnsC molecules (Figures 5A and 5B). The TnsC heptamer ring has a convex and a concave side, with the N-terminal convex side binding to TniQ (Figure 5A). Each TnsC rotates around the helix axis by ~49.3° and rises by ~1.9 Å. This geometry explains why TnsC cannot form a filament because an eighth TnsC would clash with the first one. DNA is bound to the positively charged central pore through sequence-independent interactions, reminiscent of DNA binding to TnsC in the type I-F and type V-K systems (Figures S6F and S6G).

TnsC adopts an overall fold of a typical AAA+ domain, with a large nucleotide-binding domain and a small helical bundle domain that is followed by a helical tail (residues 311–349) (Figure 6A). Residues 350–383 at the very end of the C terminus were not observed in the cryo-EM map, indicating that this region is flexible. However, the deletion of the flexible region abolished transposition activity, suggesting that this region may play an important role possibly in the recruitment of TnsAB (Figure S1J).[40] All seven TnsC molecules are in an ATP-bound state, with the ATP molecules located between neighbor subunits (Figure 6B). The *in vitro* transposition activity was greatly reduced

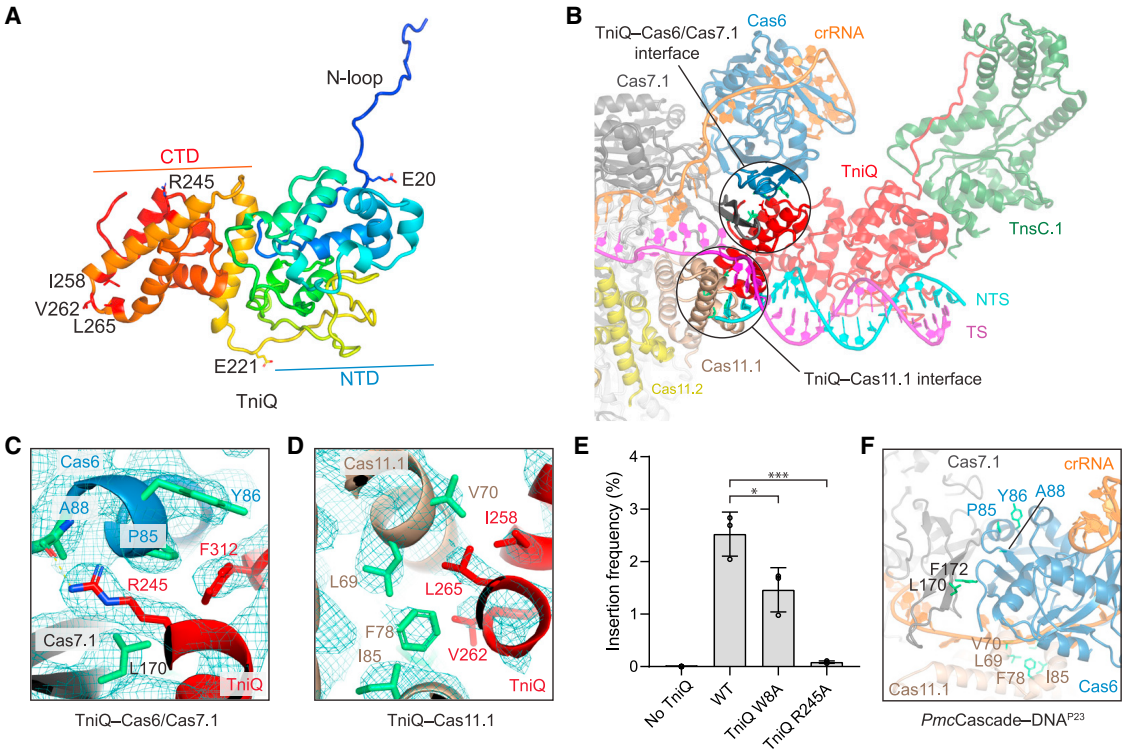


Figure 4. Structure of TniQ and *Pmc*Cascade-TniQ interactions

(A) Atomic model of TniQ colored in rainbow mode.

(B) Two binding sites between TniQ and *Pmc*Cascade in the context of *Pmc*Cascade-DNA-TniQ-TnsC.

(C) Close-up view of the interactions between TniQ and Cas6/Cas7.1. Cryo-EM densities are shown in mesh.

(D) Close-up view of the interactions between TniQ and Cas11.1. Cryo-EM densities are shown in mesh.

(E) qPCR analysis results for *in vitro* transposition assay using TniQ variants. Data are presented as mean values ± SD (n = 3). Two-sided t test was performed (*p < 0.05; ***p < 0.001).

(F) Structural mapping of key residues (colored in green) required for TniQ association in *Pmc*Cascade-DNA^{P23}. See also Figures S3 and S6.

when ATP was substituted with a non-hydrolyzable ATP analog (AMP-PNP) or when TnsC was replaced with an ATP-hydrolysis-deficient TnsC, TnsC^{E172Q} (Figure S6H). These findings indicate that ATP hydrolysis in *Pmc*CAST plays a crucial role in transposition activity. Additionally, when ADP was used instead of ATP, TnsC failed to form heptamers (Figure S4C), indicating that TnsC heptamer formation is dependent on ATP binding. This is consistent with findings on TnsC in the prototype Tn7 transposon.[41]

TniQ binds to the seam region with direct contact with both TnsC.1 and TnsC.7 (Figures 5A and 5B), with TnsC.1 being the major binding partner. The N-terminal loop of TniQ (residues 1–20) makes extensive contacts with TnsC.1 (Figures 6C and S4I). Particularly, W8 of TniQ packs against R69 of TnsC, establishing a pi-pi interaction. Consistently, the substitution of W8 with alanine (W8A) significantly reduced the RNA-guided transposition activity (Figure 4E). When the N-loop of TniQ was deleted, the co-purification of TnsC-TniQ was not successful, suggesting that the N-loop is critical for forming a stable TnsC-TniQ complex. In addition to the N-loop, residues 21–24 and 46–59 of TniQ make contacts with the small AAA+ domains of TnsC.1 through their complementary surfaces (Figure 6D). Modeling of TniQ binding to TnsC.2–7 shows a severe clash between TniQ and the previous TnsC molecule (Figures 6E and 6F). This analysis explains why TniQ is only bound to TnsC.1, because the gap between TnsC.1 and TnsC.7 prevents a clash between TniQ and TnsC.

We also determined the cryo-EM structures of the TnsC heptamer alone at 3.0 Å and the TnsC-DNA complex at 2.9 Å, both in ATP-bound states (Figures S4B, S4C, S4G, and S4H; Table S3). In both structures, TnsC.2–6 adopt structures similar to those in TnsC-TniQ-DNA. However, the densities corresponding to TnsC.1 and TnsC.7 in the TnsC heptamer alone or the TnsC-DNA complex are weak, indicating that TnsC.1 and TnsC.7 may be conformationally flexible, dissociated in a subset of particles, or both (Figures S6I and S6J). Therefore, TniQ binding likely stabilizes TnsC.1 and TnsC.7 at the seam region. Moreover, TniQ binding also promotes DNA bending, which will be discussed further below.

Target DNA recognition

Ultimately, we analyzed how the target DNA is recognized by CAST components, from the PAM sequence to near the insertion site in the structure of the *Pmc*Cascade-DNA-TniQ-TnsC recruitment complex. The PAM duplex is positioned at the interface between Cas8 and Cas5, with both subunits contributing residues that directly interact with the ATG PAM sequence (Figures 7A and 7B). In particular, K78 from Cas5 establishes a hydrogen

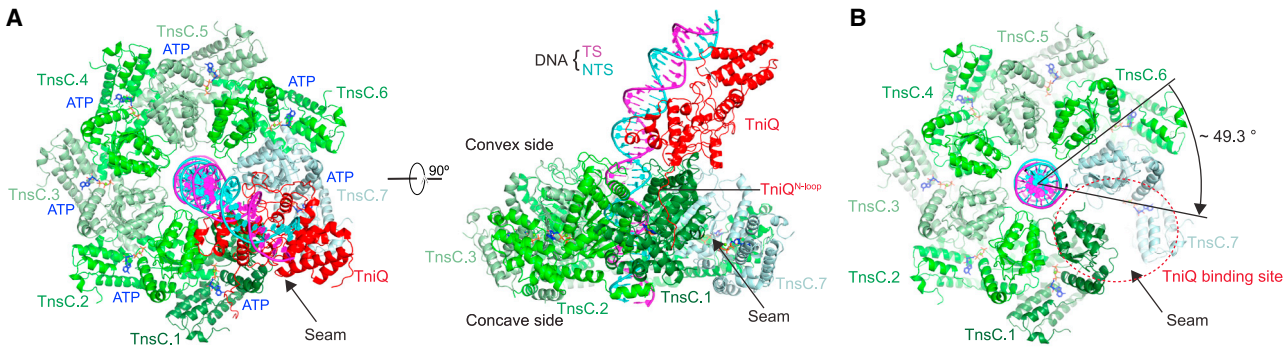


Figure 5. Overall structure of TnsC-TniQ-DNA

(A) Atomic model of TnsC-TniQ-DNA shown in top and side views.

(B) Structure of the TnsC heptamer in TnsC-TniQ-DNA, highlighting the seam between TnsC.1 and TnsC.7. See also Figures S4 and S6.

bond with the base of A(−2) in the TS through the major groove (Figures 7A and 7C). Through the minor groove, the carboxamide group of N117 from Cas8 forms a hydrogen bond with the base of T(−2) in the NTS, whereas K118 from Cas8 interacts with A(−3) in the NTS. Additionally, several polar or positively charged residues recognize the PAM duplex through the phosphate backbones, such as R76 from Cas5 (Figures 7A and 7B). To verify the structural observations, we mutated the residues responsible for PAM recognition (e.g., R76A and K78A of Cas5). Single alanine substitution of either R76 or K78 of Cas5 significantly decreased *in vitro* transposition activity as measured by quantitative PCR, whereas other residues were tolerant to alanine substitutions (Figures 7D and S7A). Electrophoresis mobility shift assay (EMSA) data revealed a substantial reduction in binding affinity for the Cas5-K78A and Cas5-R76A mutations compared with the wild-type Cascade, suggesting that the diminished transposition activity of these mutants may be attributed to their reduced affinity for the target DNA (Figures 7E and S7B).

The TS (positions 12–35) of the target DNA is held by the Cas11/Cas8$^{C\text{-}lobe}$ filament through its positively charged groove (Figure 7F). Two distinct interaction modes were observed in the *Pmc*Cascade-DNA^{P23} and *Pmc*Cascade-DNA-TniQ-TnsC states, as described above (Figure 3F). The NTS is held by Cas11 molecules in *Pmc*Cascade, similar to Cascade in type I-C[42] and I-D[43] systems; however, the relatively weak density corresponding to the NTS does not allow for an accurate model building of this region (Figure 2B). The PAM-distal end duplex (positions 36–75) is bound to TniQ and the TnsC heptamer in a sequence-independent manner primarily through electrostatic interactions (Figure 7G). Remarkably, the DNA spanning positions 36–75 is bent by approximately 50° around TniQ, as measured by the Curves+ program[44] (Figure S7C). In *Pmc*Cascade-DNA-TniQ-TnsC, the C termini of TnsC are located at the farthest end from the PAM (position 75) and are likely involved in recruiting TnsAB to catalyze DNA insertion (Figure 7G).

DISCUSSION

In this study, we used biochemical reconstitution and determined the structure of *Pmc*Cascade-DNA-TniQ-TnsC to understand the molecular mechanism of Tn7 recruitment to a type I-B2 Cascade. Our findings not only reveal mechanistic similarities between type I-B2 and V-K CAST systems but also highlight mechanistic diversity.

In both type V-K *Sh*CAST and type I-B2 *Pmc*CAST, conserved proteins such as TnsB, TnsC, and TniQ are required for RNA-guided transposition (Table S4). TniQ is recruited to the CRISPR effectors first in this process. However, the mechanistic details differ between the two systems. In *Sh*CAST, TniQ is recruited to Cas12k through its interaction with a large guide RNA and a host factor, the ribosomal protein S15.[16] In contrast, TniQ recruitment in *Pmc*CAST involves multiple Cas subunits in the *Pmc*Cascade, including Cas6, Cas7, and Cas11. Interestingly, our analysis of the *Pmc*CAST recruitment complex did not identify any host factors, although the possibility of a host factor being involved in the *in vivo* process cannot be completely ruled out.

The oligomerization state of TnsC and the association between TnsC and TniQ also show variations. In *Sh*CAST, TniQ is bound to a TnsC hexamer, which can extend to form a TnsC filament[12,13,16,17]; TniQ can bind to two sites on one end of the TnsC filament, but only one TniQ participates in the assembly of the holo transpososome.[13,17] However, in *Pmc*CAST, TnsC forms a heptamer, which appears to be a more common oligomeric state of TnsC in Tn7-like transposons.[22,41] Our structural analysis shows that the geometry of TnsC in *Pmc*CAST does not allow it to form a filament, and only one TniQ is able to bind to the seam region of the TnsC heptamer.

*Sh*CAST lacks the TnsA component that is present in *Pmc*CAST, which leads to differences in their functions. Although *Sh*CAST mediates cointegration, *Pmc*CAST can catalyze simple insertions. However, cointegration can be converted into simple insertions in an *in vivo* environment through a process involving the site-specific recombinase.[45] It will be interesting to investigate how TnsC in *Pmc*CAST recruits the TnsAB transposase in future studies.

Although the type I-F *Vch*CAST system shares a few similar features with the type I-B2 *Pmc*CAST system, such as utilizing a multiple-subunit Cascade effector for target DNA recognition and a TnsC heptamer, there are significant differences between these two systems. First, *Pmc*Cascade contains Cas11 subunits, which are critical for the RNA-guided transposition activity but are not

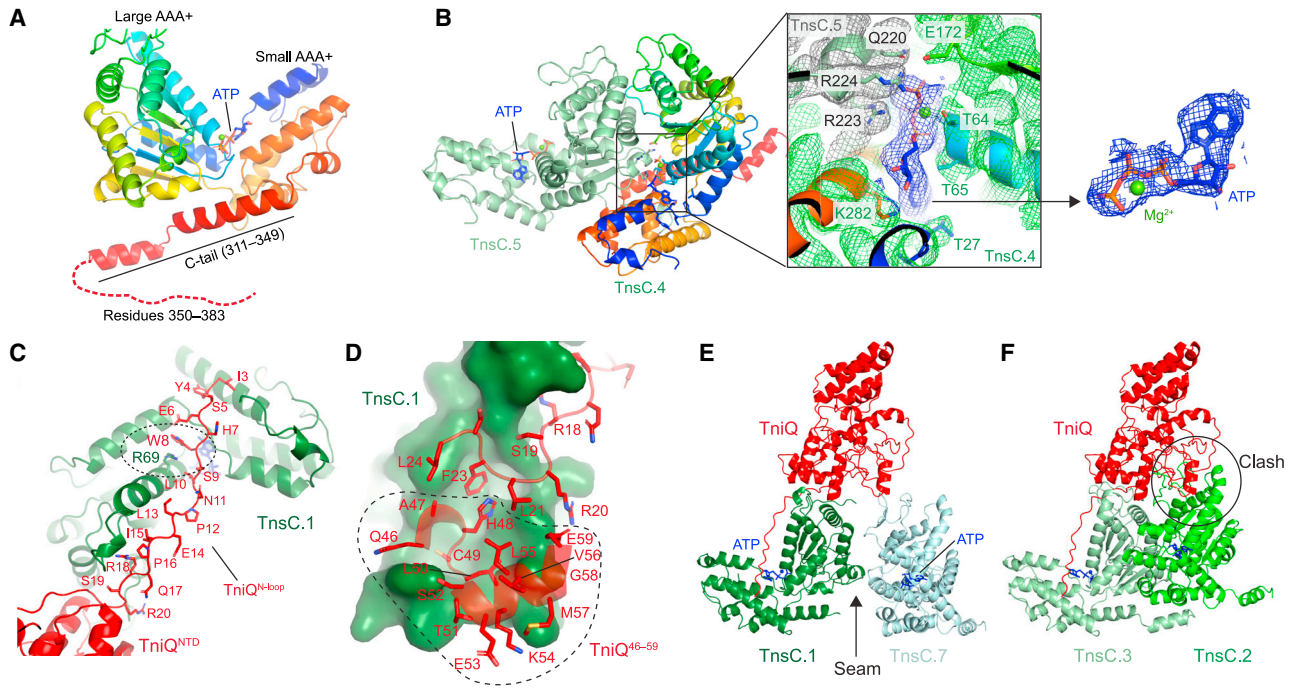


Figure 6. TnsC structure and TnsC-TniQ interactions

(A) Atomic model of TnsC isolated from the TnsC-TniQ-DNA structure.

(B) Structure of two neighbor TnsC molecules in TnsC-TniQ-DNA, highlighting the ATP molecule bound in between them.

(C) Close-up view showing the interactions between the N-loop of TniQ and TnsC.1.

(D) Close-up view showing the interactions between the NTD of TniQ (residues 18–24 and 46–59) and TnsC.1.

(E) TniQ is located at the seam between TnsC.1 and TnsC.7.

(F) Structural modeling of TniQ to TnsC.3 shows severe clash between TniQ and TnsC.2. See also Figures S4 and S6.

present in *Vch*Cascade. Second, in *Vch*CAST, a TniQ dimer directly binds to *Vch*Cascade prior to target DNA binding,[18–21] whereas in *Pmc*CAST, target DNA recognition and full R-loop formation are required for TniQ recruitment. Third, although the TnsC proteins in both systems form heptamers, the seam observed in *Pmc*TnsC is not observed in *Vch*TnsC. The seam region plays a crucial role in TniQ association in *Pmc*CAST, but the structure of TnsC-TniQ is not yet available for *Vch*CAST.

In summary, our results advance our structural and mechanistic understanding of the RNA-guided DNA transposition in CASTs by providing insights into the type I-B2 *Pmc*CAST. In particular, our results demonstrate that the mechanism for Tn7-like transposon recruitment by *Pmc*Cascade is distinct from those reported in the type V-K and I-F systems. The diverse mechanisms are consistent with a phylogenetic analysis, which suggests that the association between CRISPR-Cas systems and Tn7-like transposons occurred independently on several occasions.[1] Recent applications of type V-K[46] and I-F[47] CASTs for gene knockin in mammalian cells have shown great promise, and it would be thrilling to explore the potential of *Pmc*CAST to further enrich the knockin toolkit. By biochemically reconstructing *Pmc*CAST and determining the *Pmc*Cascade-DNA-TniQ-TnsC structure, we have established a foundation for future mechanistic studies of type I CASTs. These studies will support the use of CASTs as high-fidelity gene knockin tools in biomedical research.

Limitations of the study

It will be important to design and conduct cellular experiments informed by the findings in this study, particularly our insights into the role of Cas11. Our methods primarily focused on testing on-target RNA-guided DNA insertions but did not address potential off-target effects in *Pmc*CAST. Subsequently, it will be vital to explore how TnsAB is recruited and activated by the Cascade-DNA-TniQ-TnsC complex. Furthermore, unraveling the mechanism of how TnsD recognizes the homing site and recruits TnsC will be crucial for achieving a more complete understanding of the biological processes involved in CASTs.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS

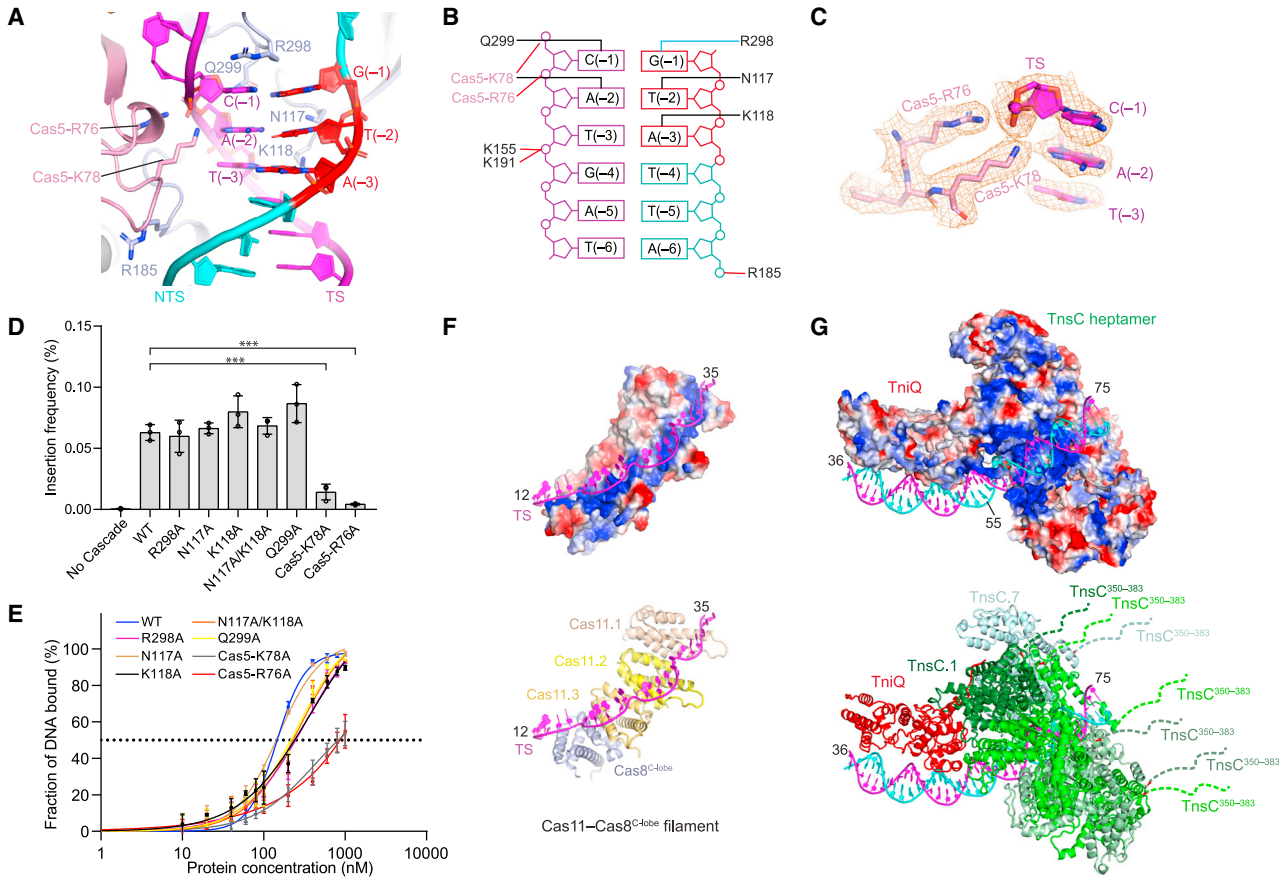


Figure 7. Target DNA recognition by components in *Pmc*CAST

(A) Interaction between *Pmc*Cascade and the PAM duplex (the ATG PAM sequence is colored in red).

(B) Schematic of interactions between *Pmc*Cascade and the PAM duplex. Black, red, and blue lines indicate hydrogen bonding, electrostatic, and stacking interactions, respectively.

(C) Cryo-EM map of *Pmc*Cascade-DNA^{P23} (in mesh) illustrating Cas5-R76 and Cas5-K78 interactions with the TS of target DNA.

(D) qPCR analysis for *in vitro* transposition assays using wild-type and mutant *Pmc*Cascade. Mutations are on Cas8 unless otherwise specified. Data are presented as mean values ± SD (n = 3). Two-sided t test was performed (***p < 0.001).

(E) EMSA quantification of wild-type and mutant Cascade binding to the target DNA. Data are presented as mean values ± SD (n = 2–4).

(F) Electrostatic interactions between the Cas11/Cas8$^{C\text{-}lobe}$ filament (shown as electrostatic potential surface and cartoon, respectively) and the TS (positions 12–35).

(G) Electrostatic interactions between TnsC-TniQ and the target DNA at the PAM-distal end (positions 36–75). TnsC-TniQ is shown as electrostatic potential surface (top) and cartoon (bottom), respectively. See also Figure S7.

- ○ Plasmid construction
- ○ Protein expression and purification
- ○ *In vitro* transposition assay
- ○ Mass spectrometry
- ○ *E. coli* transposition assay
- ○ Polymerase Chain Reaction (PCR)
- ○ Quantitative PCR
- ○ Electrophoretic mobility shift assay (EMSA)
- ○ Preparation of the *Pmc*Cascade–DNA complex
- ○ Preparation of the TnsC oligomer
- ○ Preparation of TnsC–DNA and TnsC–TniQ–DNA
- ○ Preparation of *Pmc*Cascade–TnsC–TniQ–DNA
- ○ Electron microscopy
- ○ Image processing
- ○ Model building and refinement
- ○ Sequence alignment
- ○ Structural visualization

● QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2023.07.010.

ACKNOWLEDGMENTS

We appreciate Phoebe A. Rice for her critical reading and insightful suggestions. We are grateful to Mark C. Hall for his valuable insights into the mass spectrometry results and helpful discussions. We thank Steven Wilson for computation, Renjian Xiao for help with TnsAB purification, and Drs. Chaylen Andolino and Uma Aryal for help with LC-MS/MS analysis. We are grateful to

Emily C. Dykhuizen and Guanming Jiao for help with qPCR assays and Robert V. Stahelin for providing access to the GE Amersham Imager 600 system. This work made use of the Purdue Cryo-EM Facility. All the LC-MS/MS experiments were performed at the Purdue Proteomics Facility, Bindley Bioscience Center. This work was supported by the NIH grant R01GM138675 to L.C.

AUTHOR CONTRIBUTIONS

L.C. supervised the study. S.W. prepared all samples and performed all biochemical assays. R.S. helped with sample preparations. S.W., T.K., and L.C. collected cryo-EM data. L.C. and S.W. processed cryo-EM data and built the structural models. S.W. and L.C. prepared the manuscript with input from C.G., R.S., and T.K.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 21, 2022
Revised: April 23, 2023
Accepted: July 10, 2023
Published: August 8, 2023

REFERENCES

1. Peters, J.E., Makarova, K.S., Shmakov, S., and Koonin, E.V. (2017). Recruitment of CRISPR-Cas systems by Tn7-like transposons. Proc. Natl. Acad. Sci. USA *114*. E7358–E7366. https://doi.org/10.1073/pnas.1709035114.
2. Strecker, J., Ladha, A., Gardner, Z., Schmid-Burgk, J.L., Makarova, K.S., Koonin, E.V., and Zhang, F. (2019). RNA-guided DNA insertion with CRISPR-associated transposases. Science *365*, 48–53. https://doi.org/10.1126/science.aax9181.
3. Klompe, S.E., Vo, P.L.H., Halpin-Healy, T.S., and Sternberg, S.H. (2019). Transposon-encoded CRISPR-Cas systems direct RNA-guided DNA integration. Nature *571*, 219–225. https://doi.org/10.1038/s41586-019-1323-z.
4. Saito, M., Ladha, A., Strecker, J., Faure, G., Neumann, E., Altae-Tran, H., Macrae, R.K., and Zhang, F. (2021). Dual modes of CRISPR-associated transposon homing. Cell *184*, 2441–2453.e18. https://doi.org/10.1016/j.cell.2021.03.006.
5. Peters, J.E., and Craig, N.L. (2001). Tn7: smarter than we thought. Nat. Rev. Mol. Cell Biol. *2*, 806–814. https://doi.org/10.1038/35099006.
6. May, E.W., and Craig, N.L. (1996). Switching from cut-and-paste to replicative Tn7 transposition. Science *272*, 401–404. https://doi.org/10.1126/science.272.5260.401.
7. Rice, P.A., Craig, N.L., and Dyda, F. (2020). Comment on "RNA-guided DNA insertion with CRISPR-associated transposases". Science *368*. https://doi.org/10.1126/science.abb2022.
8. Strecker, J., Ladha, A., Makarova, K.S., Koonin, E.V., and Zhang, F. (2020). Response to comment on "RNA-guided DNA insertion with CRISPR-associated transposases". Science *368*. https://doi.org/10.1126/science.abb2920.
9. Bainton, R.J., Kubo, K.M., Feng, J.N., and Craig, N.L. (1993). Tn7 transposition: target DNA recognition is mediated by multiple Tn7-encoded proteins in a purified in vitro system. Cell *72*, 931–943. https://doi.org/10.1016/0092-8674(93)90581-a.
10. Shi, Q., Straus, M.R., Caron, J.J., Wang, H., Chung, Y.S., Guarné, A., and Peters, J.E. (2015). Conformational toggling controls target site choice for the heteromeric transposase element Tn7. Nucleic Acids Res. *43*, 10734–10745. https://doi.org/10.1093/nar/gkv913.
11. Xiao, R., Wang, S., Han, R., Li, Z., Gabel, C., Mukherjee, I.A., and Chang, L. (2021). Structural basis of target DNA recognition by CRISPR-Cas12k for RNA-guided DNA transposition. Mol. Cell *81*, 4457–4466.e5. https://doi.org/10.1016/j.molcel.2021.07.043.
12. Querques, I., Schmitz, M., Oberli, S., Chanez, C., and Jinek, M. (2021). Target site selection and remodelling by type V CRISPR-transposon systems. Nature *599*, 497–502. https://doi.org/10.1038/s41586-021-04030-z.
13. Park, J.U., Tsai, A.W., Mehrotra, E., Petassi, M.T., Hsieh, S.C., Ke, A., Peters, J.E., and Kellogg, E.H. (2021). Structural basis for target site selection in RNA-guided DNA transposition systems. Science *373*, 768–774. https://doi.org/10.1126/science.abi8976.
14. Park, J.U., Tsai, A.W., Chen, T.H., Peters, J.E., and Kellogg, E.H. (2022). Mechanistic details of CRISPR-associated transposon recruitment and integration revealed by cryo-EM. Proc. Natl. Acad. Sci. USA *119*, e2202590119. https://doi.org/10.1073/pnas.2202590119.
15. Tenjo-Castaño, F., Sofos, N., López-Méndez, B., Stutzke, L.S., Fuglsang, A., Stella, S., and Montoya, G. (2022). Structure of the TnsB transposase-DNA complex of type V-K CRISPR-associated transposon. Nat. Commun. *13*, 5792. https://doi.org/10.1038/s41467-022-33504-5.
16. Schmitz, M., Querques, I., Oberli, S., Chanez, C., and Jinek, M. (2022). Structural basis for the assembly of the type V CRISPR-associated transposon complex. Cell *185*, 4999–5010.e17. https://doi.org/10.1016/j.cell.2022.11.009.
17. Park, J.U., Tsai, A.W., Rizo, A.N., Truong, V.H., Wellner, T.X., Schargel, R.D., and Kellogg, E.H. (2023). Structures of the holo CRISPR RNA-guided transposon integration complex. Nature *613*, 775–782. https://doi.org/10.1038/s41586-022-05573-5.
18. Halpin-Healy, T.S., Klompe, S.E., Sternberg, S.H., and Fernández, I.S. (2020). Structural basis of DNA targeting by a transposon-encoded CRISPR-Cas system. Nature *577*, 271–274. https://doi.org/10.1038/s41586-019-1849-0.
19. Li, Z., Zhang, H., Xiao, R., and Chang, L. (2020). Cryo-EM structure of a type I-F CRISPR RNA guided surveillance complex bound to transposition protein TniQ. Cell Res. *30*, 179–181. https://doi.org/10.1038/s41422-019-0268-y.
20. Jia, N., Xie, W., de la Cruz, M.J., Eng, E.T., and Patel, D.J. (2020). Structure-function insights into the initial step of DNA integration by a CRISPR-Cas-Transposon complex. Cell Res. *30*, 182–184. https://doi.org/10.1038/s41422-019-0272-2.
21. Wang, B., Xu, W., and Yang, H. (2020). Structural basis of a Tn7-like transposase recruitment and DNA loading to CRISPR-Cas surveillance complex. Cell Res. *30*, 185–187. https://doi.org/10.1038/s41422-020-0274-0.
22. Hoffmann, F.T., Kim, M., Beh, L.Y., Wang, J., Vo, P.L.H., Gelsinger, D.R., George, J.T., Acree, C., Mohabir, J.T., Fernández, I.S., et al. (2022). Selective TnsC recruitment enhances the fidelity of RNA-guided transposition. Nature *609*, 384–393. https://doi.org/10.1038/s41586-022-05059-4.
23. Petassi, M.T., Hsieh, S.C., and Peters, J.E. (2020). Guide RNA categorization enables target site choice in Tn7-CRISPR-Cas transposons. Cell *183*, 1757–1771.e18. https://doi.org/10.1016/j.cell.2020.11.005.
24. Salis, H.M., Mirsky, E.A., and Voigt, C.A. (2009). Automated design of synthetic ribosome binding sites to control protein expression. Nat. Biotechnol. *27*, 946–950. https://doi.org/10.1038/nbt.1568.
25. McBride, T.M., Schwartz, E.A., Kumar, A., Taylor, D.W., Fineran, P.C., and Fagerlund, R.D. (2020). Diverse CRISPR-Cas complexes require independent translation of small and large subunits from a single gene. Mol. Cell *80*, 971–979.e7. https://doi.org/10.1016/j.molcel.2020.11.003.
26. Tan, R., Krueger, R.K., Gramelspacher, M.J., Zhou, X., Xiao, Y., Ke, A., Hou, Z., and Zhang, Y. (2022). Cas11 enables genome engineering in human cells with compact CRISPR-Cas3 systems. Mol. Cell *82*, 852–867.e5. https://doi.org/10.1016/j.molcel.2021.12.032.
27. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.
28. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. Nat. Methods *19*, 679–682. https://doi.org/10.1038/s41592-022-01488-1.

29. Ronning, D.R., Li, Y., Perez, Z.N., Ross, P.D., Hickman, A.B., Craig, N.L., and Dyda, F. (2004). The carboxy-terminal portion of TnsC activates the Tn7 transposase through a specific interaction with TnsA. EMBO J. *23*, 2972–2981. https://doi.org/10.1038/sj.emboj.7600311.

30. Montaño, S.P., Pigli, Y.Z., and Rice, P.A. (2012). The μ transpososome structure sheds light on DDE recombinase evolution. Nature *491*, 413–417. https://doi.org/10.1038/nature11602.

31. Söding, J., Biegert, A., and Lupas, A.N. (2005). The HHpred interactive server for protein homology detection and structure prediction. Nucleic Acids Res. *33*. W244–W248. https://doi.org/10.1093/nar/gki408.

32. Jackson, R.N., Golden, S.M., van Erp, P.B., Carter, J., Westra, E.R., Brouns, S.J., van der Oost, J., Terwilliger, T.C., Read, R.J., and Wiedenheft, B. (2014). Structural biology. Crystal structure of the CRISPR RNA-guided surveillance complex from Escherichia coli. Science *345*, 1473–1479. https://doi.org/10.1126/science.1256328.

33. Zhao, H.T., Sheng, G., Wang, J.Y., Wang, M., Bunkoczi, G., Gong, W.M., Wei, Z.Y., and Wang, Y.L. (2014). Crystal structure of the RNA-guided immune surveillance Cascade complex in Escherichia coli. Nature *515*, 147–150. https://doi.org/10.1038/nature13733.

34. Pausch, P., Müller-Esparza, H., Gleditzsch, D., Altegoer, F., Randau, L., and Bange, G. (2017). Structural variation of type I-F CRISPR RNA guided DNA surveillance. Mol. Cell *67*, 622–632.e4. https://doi.org/10.1016/j.molcel.2017.06.036.

35. Mulepati, S., Héroux, A., and Bailey, S. (2014). Structural biology. Crystal structure of a CRISPR RNA-guided surveillance complex bound to a ssDNA target. Science *345*, 1479–1484. https://doi.org/10.1126/science.1256996.

36. Guo, T.W., Bartesaghi, A., Yang, H., Falconieri, V., Rao, P., Merk, A., Eng, E.T., Raczkowski, A.M., Fox, T., Earl, L.A., et al. (2017). Cryo-EM structures reveal mechanism and inhibition of DNA targeting by a CRISPR-Cas surveillance complex. Cell *171*, 414–426.e12. https://doi.org/10.1016/j.cell.2017.09.006.

37. Wiedenheft, B., van Duijn, E., Bultema, J.B., Waghmare, S.P., Zhou, K., Barendregt, A., Westphal, W., Heck, A.J., Boekema, E.J., Dickman, M.J., et al. (2011). RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. Proc. Natl. Acad. Sci. USA *108*, 10092–10097. https://doi.org/10.1073/pnas.1102716108.

38. Semenova, E., Jore, M.M., Datsenko, K.A., Semenova, A., Westra, E.R., Wanner, B., van der Oost, J., Brouns, S.J., and Severinov, K. (2011). Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. Proc. Natl. Acad. Sci. USA *108*, 10098–10103. https://doi.org/10.1073/pnas.1104144108.

39. Holm, L. (2022). Dali server: structural unification of protein families. Nucleic Acids Res. *50*. W210–W215. https://doi.org/10.1093/nar/gkac387.

40. Skelding, Z., Queen-Baker, J., and Craig, N.L. (2003). Alternative interactions between the Tn7 transposase and the Tn7 target DNA binding protein regulate target immunity and transposition. EMBO J. *22*, 5904–5917. https://doi.org/10.1093/emboj/cdg551.

41. Shen, Y., Gomez-Blanco, J., Petassi, M.T., Peters, J.E., Ortega, J., and Guarné, A. (2022). Structural basis for DNA targeting by the Tn7 transposon. Nat. Struct. Mol. Biol. *29*, 143–151. https://doi.org/10.1038/s41594-022-00724-8.

42. O'Brien, R.E., Santos, I.C., Wrapp, D., Bravo, J.P.K., Schwartz, E.A., Brodbelt, J.S., and Taylor, D.W. (2020). Structural basis for assembly of non-canonical small subunits into type I-C Cascade. Nat. Commun. *11*, 5931. https://doi.org/10.1038/s41467-020-19785-8.

43. Schwartz, E.A., McBride, T.M., Bravo, J.P.K., Wrapp, D., Fineran, P.C., Fagerlund, R.D., and Taylor, D.W. (2022). Structural rearrangements allow nucleic acid discrimination by type I-D Cascade. Nat. Commun. *13*, 2829. https://doi.org/10.1038/s41467-022-30402-8.

44. Lavery, R., Moakher, M., Maddocks, J.H., Petkeviciute, D., and Zakrzewska, K. (2009). Conformational analysis of nucleic acids revisited: curves+. Nucleic Acids Res. *37*, 5917–5929. https://doi.org/10.1093/nar/gkp608.

45. Kholodii, G.Y., Mindlin, S.Z., Bass, I.A., Yurieva, O.V., Minakhina, S.V., and Nikiforov, V.G. (1995). Four genes, two ends, and a res region are involved in transposition of Tn5053: a paradigm for a novel family of transposons carrying either a mer operon or an integron. Mol. Microbiol. *17*, 1189–1200. https://doi.org/10.1111/j.1365-2958.1995.mmi_17061189.x.

46. Tou, C.J., Orr, B., and Kleinstiver, B.P. (2023). Precise cut-and-paste DNA insertion using engineered type V-K CRISPR-associated transposases. Nat. Biotechnol. *2023*. https://doi.org/10.1038/s41587-022-01574-x.

47. Lampe, G.D., King, R.T., Halpin-Healy, T.S., Klompe, S.E., Hogan, M.I., Vo, P.L.H., Tang, S., Chavez, A., and Sternberg, S.H. (2023). Targeted DNA integration in human cells without double-strand breaks using CRISPR-associated transposases. Nat. Biotechnol. *2023*. https://doi.org/10.1038/s41587-023-01748-1.

48. Lau, Y.K., Baytshtok, V., Howard, T.A., Fiala, B.M., Johnson, J.M., Carter, L.P., Baker, D., Lima, C.D., and Bahl, C.D. (2018). Discovery and engineering of enhanced SUMO protease enzymes. J. Biol. Chem. *293*, 13224–13233. https://doi.org/10.1074/jbc.RA118.004146.

49. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., and Ferrin, T.E. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. J. Comput. Chem. *25*, 1605–1612. https://doi.org/10.1002/jcc.20084.

50. Pettersen, E.F., Goddard, T.D., Huang, C.C., Meng, E.C., Couch, G.S., Croll, T.I., Morris, J.H., and Ferrin, T.E. (2021). UCSF ChimeraX: structure visualization for researchers, educators, and developers. Protein Sci. *30*, 70–82. https://doi.org/10.1002/pro.3943.

51. Afonine, P.V., Poon, B.K., Read, R.J., Sobolev, O.V., Terwilliger, T.C., Urzhumtsev, A., and Adams, P.D. (2018). Real-space refinement in PHENIX for cryo-EM and crystallography. Acta Crystallogr. D Struct. Biol. *74*, 531–544. https://doi.org/10.1107/S2059798318006551.

52. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. *66*, 486–501. https://doi.org/10.1107/S0907444910007493.

53. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat. Methods *14*, 290–296. https://doi.org/10.1038/nmeth.4169.

54. Robert, X., and Gouet, P. (2014). Deciphering key features in protein structures with the new ENDscript server. Nucleic Acids Res. *42*. W320–W324. https://doi.org/10.1093/nar/gku316.

55. Stothard, P. (2000). The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. BioTechniques *28*, 1102–1104. https://doi.org/10.2144/00286ir01.

56. Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nat. Methods *9*, 671–675.

57. Cox, J., and Mann, M. (2008). MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat. Biotechnol. *26*, 1367–1372. https://doi.org/10.1038/nbt.1511.

58. Biesiada, M., Purzycka, K.J., Szachniuk, M., Blazewicz, J., and Adamiak, R.W. (2016). Automated RNA 3D structure prediction with RNAComposer. Methods Mol. Biol. *1490*, 199–215. https://doi.org/10.1007/978-1-4939-6433-8_13.

59. McWilliam, H., Li, W., Uludag, M., Squizzato, S., Park, Y.M., Buso, N., Cowley, A.P., and Lopez, R. (2013). Analysis tool web services from the EMBL-EBI. Nucleic Acids Res. *41*. W597–W600. https://doi.org/10.1093/nar/gkt376.

60. Kim, S.Q., Mohallem, R., Franco, J., Buhman, K.K., Kim, K.H., and Aryal, U.K. (2022). Global landscape of protein complexes in postprandial-state livers from diet-induced obese and lean mice. Biochem. Biophys. Res. Commun. *629*, 40–46. https://doi.org/10.1016/j.bbrc.2022.08.070.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
One Shot™ TOP10 Chemically Competent E. coli	Thermo Fisher Scientific	Cat#C404010
One Shot PIR1 Chemically Competent E. coli	Thermo Fisher Scientific	Cat#C101010
BL21-CodonPlus (DE3)-RIPL Competent Cells	Agilent	Cat#230280
Rosetta (DE3) pLysS Competent cells	Novagen	Cat#70956
BL21(DE3) Competent Cells	Thermo Fisher Scientific	Cat#EC0114
Chemicals, peptides, and recombinant proteins		
Thermo Scientific™ Phusion High-Fidelity PCR Master Mix with HF Buffer	Thermo Fisher Scientific	Cat#F531L
PrimeSTAR® Max DNA Polymerase	Takara	Cat#R045B
Thermo Scientific™ Phusion Hot Start II DNA Polymerase (2 U/μL)	Thermo Fisher Scientific	Cat#F549L
DpnI	New England Biolabs	Cat#R0176L
BamHI-HF	New England Biolabs	Cat#R3136S
QIAquick Gel Extraction Kit	Qiagen	Cat#28706
QIAprep Spin Miniprep Kit	Qiagen	Cat#27106
QIAquick PCR Purification Kit	Qiagen	Cat#28106
In-Fusion® HD Cloning Kit	Takara	Cat#639650
Terrific Broth	Thermo Fisher Scientific	Cat#BP97285
Isopropyl b-D-thiogalactopyranoside (IPTG)	Thermo Fisher Scientific	Cat#BP1755100
SUMO protease	Home-made	N/A
Ni-NTA agarose resin	Qiagen	Cat#30430
Strep-Tactin Superflow Plus	Qiagen	Cat#30004
Pierce Protease Inhibitor Tablets	Thermo Fisher Scientific	Cat#PIA32963
TaqMan Fast Advanced Master Mix	Thermo Fisher Scientific	Cat#44-445-57
Ampicillin sodium salt	Thermo Fisher Scientific	Cat#BP1760-25
Chloramphenical, USP grade	GoldBio	Cat#C-105-100
Kanamycin Monosulfate, USP grade	GoldBio	Cat#K-120-100
Spectinomycin Dihydrochloride Pentahydrate, USP Grade	GoldBio	Cat#S-140-50
Adenosine triphosphate (ATP)	Thermo Fisher Scientific	Cat#BP413-25
Adenosine diphosphate (ADP)	Research Products International	Cat#A11220-10.0
Adenylyl-imidodiphosphate (AMP-PNP)	Sigma	Cat#10102547001
Critical commercial assays		
HiTrap Q HP	Cytiva	Cat#17115401
Superdex 200 Increase 10/300 column	Cytiva	Cat#28990944
Deposited data		
PmcCascade–DNAP23 coordinate	This study	PDB: 8FCJ
PmcCascade–DNAP23 map	This study	EMDB: EMD-28980
PmcCascade–DNA–TniQ–TnsC coordinate	This study	PDB: 8FCU
*Pmc*Cascade–DNA–TniQ–TnsC map	This study	EMDB: EMD-28993

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
*Pmc*Cascade–DNA–TniQ–TnsC composite coordinate	This study	PDB: 8FF4
*Pmc*Cascade–DNA–TniQ–TnsC composite map	This study	EMDB: EMD-29039
*Pmc*Cascade coordinate	This study	PDB: 8FD2
*Pmc*Cascade map	This study	EMDB: EMD-29000
*Pmc*Cascade–PAM coordinate	This study	PDB: 8FD3
*Pmc*Cascade–PAM map	This study	EMDB: EMD-29001
*Pmc*Cascade–DNA$^{full\ R\text{-}loop}$ coordinate	This study	PDB: 8FF5
*Pmc*Cascade–DNA$^{full\ R\text{-}loop}$ map	This study	EMDB: EMD-29040
TnsC–TniQ–DNA coordinate	This study	PDB: 8FCV
TnsC–TniQ–DNA map	This study	EMDB: EMD-28994
TnsC–DNA coordinate	This study	PDB: 8FCW
TnsC–DNA map	This study	EMDB: EMD-28997
TnsC coordinate	This study	PDB: 8FCX
TnsC map	This study	EMDB: EMD-28998
Oligonucleotides		
Oligonucleotides	This study	Table S5
Recombinant DNA		
pTwinStrep-SUMO-PmcTnsAB	Saito et al.[4]	Addgene Plasmid #168164
pTwinStrep-SUMO-PmcTnsC	Saito et al.[4]	Addgene Plasmid #168165
pTarget(PmcCAST)	Saito et al.[4]	Addgene Plasmid #168163
pDonor(PmcCAST)	Saito et al.[4]	Addgene Plasmid #168162
pHelper(PmcCAST)_PmcPSP1_ΔTnsD	Saito et al.[4]	Addgene Plasmid #168153
pHelper(PmcCAST)_entry_ΔTnsD	Saito et al.[4]	Addgene Plasmid #168156
pHelper(PmcCAST)_PmcPSP1	Saito et al.[4]	Addgene Plasmid #168151
pET-StrepII-TEV-LIC	Scott Gradia	Addgene Plasmid # 29664
pDonor_ShCAST_kanR	Strecker et al.[2]	Addgene Plasmid #127924
pUlp1_R3_sumo_protease	Addgene[48]	Addgene Plasmid #113671
pCascade_purification	This study	N/A
pTniQ_purification	This study	N/A
pDonor_PmcCAST_kanR	This study	N/A
pCascade_purification variant mutants	This study	N/A
pTniQ_purification variant mutants	This study	N/A
pTwinStrep-SUMO-PmcTnsAB variant mutants	This study	N/A
pTwinStrep-SUMO-PmcTnsC variant mutants	This study	N/A
pHelper(PmcCAST)_PmcPSP1_ΔTnsD variant mutants	This study	N/A
pTarget(PmcCAST) variant mutants	This study	N/A
Software and algorithms		
PyMOL	Schrodinger LLC	https://pymol.org/2/
UCSF Chimera 1.12	Pettersen et al.[49]	https://www.cgl.ucsf.edu/chimera/
UCSF ChimeraX 1.2	Pettersen et al.[50]	https://www.cgl.ucsf.edu/chimerax/
Phenix 1.20.1	Afonine et al.[51]	https://phenix-online.org/
Coot 0.9.8	Emsley et al.[52]	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
CryoSPARC v3.3.2	Punjani et al.[53]	https://cryosparc.com/

(*Continued on next page*)

Continued		
REAGENT or RESOURCE	SOURCE	IDENTIFIER
ESPript 3	Robert and Gouet[54]	https://espript.ibcp.fr/ESPript/ESPript/
Sequence Manipulation Suite 2	Stothard[55]	https://www.bioinformatics.org/sms2/ident_sim.html
ImageJ	Schneider et al.[56]	https://imagej.nih.gov/ij/download.html
GraphPad Prism 8	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
Curves+	Lavery et al.[44]	http://curvesplus.bsc.es/analyse
RBS Calculator 2.1	Salis et al.[24]	https://www.denovodna.com/software/predict_rbs_calculator
MaxQuant v2.0.3.0	Cox and Mann[57]	https://www.maxquant.org/
RNAComposer	Biesiada et al.[58]	https://rnacomposer.cs.put.poznan.pl/
Clustal Omega	McWilliam et al.[59]	https://www.ebi.ac.uk/Tools/msa/clustalo/
Other		
Mendeley Data	This study	https://doi.org/10.17632/cz7frtx4np.1

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Leifu Chang (lchang18@purdue.edu).

Materials availability

Plasmids generated in this study will be made available upon request made to the lead contact.

Data and code availability

- The Electron Microscopy Data Bank (EMDB) and Protein Data Bank (PDB) depositions are publicly available as of the date of publication. Accession numbers are listed in the key resources table.
- This paper does not report original code.
- Additional Supplemental Items are available from Mendeley Data at https://doi.org/10.17632/cz7frtx4np.1. The dataset will be publicly accessible. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

The plasmid DNAs were amplified in *E. coli* Top10. Cascade and transposition proteins were expressed in *E. coli* Rosetta (DE3)pLysS and BL21-CodonPlus (DE3)-RIPL cells, respectively. Products of *in vitro* transposition were transformed into Top10 competent cells for amplification.

METHOD DETAILS

Plasmid construction

The primers used to construct plasmids are listed in Table S5. The expression plasmid for *Pmc*Cascade was modified from pHelper(PmcCAST)_*Pmc*PSP1 (Addgene: #168151). To delete the genes of TnsAB, TnsC, TnsD, and TniQ, pHelper was linearized with primers Cas_f and Cas_r. PCR product was purified by a gel extraction kit (Qiagen) and then self-ligated with the In-Fusion Cloning Kit (Takara). After sequencing confirmation, an 8xHis tag was added to the N terminus of Cas8 with primers *Pmc*_add_his_f and *Pmc*_add_his_r, and the product was self-ligated with the In-Fusion Cloning Kit.

The gene of *Pmc*TniQ was amplified from pHelper(PmcCAST)_*Pmc*PSP1 with primers *Pmc*TniQ_f and *Pmc*TniQ_r. The purified PCR product was inserted into pET-StrepII-TEV-LIC (Addgene: #29664; a gift from Scott Gradia) with the In-Fusion Cloning Kit.

Single amino acid mutations were generated by the QuickChange site-directed mutagenesis method. Mutated target plasmids with different spacer lengths were constructed by ligating inverse PCR-amplified backbone with mutations-bearing DNA oligonucleotides via the In-Fusion Cloning Kit. Mutations to target plasmids for spacer mismatch experiment were generated by the

QuickChange site-directed mutagenesis except for the 19–23 mismatch, which was made via In-Fusion Cloning Kit. All mutations were confirmed by Sanger sequencing. Primers and oligonucleotides used are listed in Table S5.

The donor plasmid containing kanamycin resistance gene was cloned by replacing the LE and RE sequences of a previous donor plasmid (Addgene #127924) for type V-K with the LE and RE sequences of type I-B2 *Pmc*CAST system.

To delete the Cas11 coding region, pHelper(PmcCAST)_*Pmc*PSP1_ΔTnsD was linearized using primers Cas11_dele_f and Cas11_dele_r. The PCR product was self-ligated with the In-Fusion Cloning Kit after gel extraction.

Protein expression and purification

*Pmc*Cascade was expressed in Rosetta (DE3) pLysS (Novagen: #70956). The expression plasmid was freshly transformed every time before *Pmc*Cascade expression. Single colonies were picked and cultured in 50 mL Lysogeny Broth with 50 μg/mL chloramphenicol and 100 μg/mL spectinomycin overnight. Cells were then inoculated into Terrific Broth and grown to 1.0 OD before induction with 0.25 mM IPTG and growth at 16 °C overnight. Cells were harvested, flash frozen, and stored at –80 °C. For purification, cell pellets were suspended in lysis buffer (50 mM Tris-HCl pH 7.5, 500 mM NaCl, 5% glycerol, 10 mM imidazole) supplemented with Pierce Protease Inhibitor Tablets (Fisher Scientific) and then disrupted by sonication. Cell lysate was clarified by centrifugation at 4 °C at 15,000 rpm for 1 h. The supernatant was filtered before loaded onto Ni-NTA resin (Qiagen). After extensive washing with lysis buffer supplemented with 30 mM imidazole, the target protein was eluted with lysis buffer supplemented with 250 mM imidazole. The eluted sample was diluted with buffer containing 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5% glycerol, 1 mM DTT and loaded onto HiTrap Q HP (Cytiva). The protein was eluted with 50 mM Tris-HCl (pH 7.5), a step-gradient of 0.1–1 M NaCl, 5% glycerol, 1 mM DTT. Fractions containing the target protein were pooled and concentrated before being loaded onto a Superdex 200 Increase column (Cytiva) with buffer containing 20 mM HEPES (pH 7.5), 200 mM NaCl, and 1 mM DTT. Fractions were concentrated and stored at –80 °C.

TnsAB and TnsC were expressed and purified similar to previous methods.[4] Specifically, both proteins were expressed in BL21-CodonPlus (DE3)-RIPL. Cells were grown to OD 0.6 in Terrific Broth with 100 μg/mL ampicillin and protein expression was induced by adding 0.5 mM IPTG followed by overnight growth at 16 °C. Cells were collected and resuspended with buffer containing 50 mM Tris-HCl (pH 8.0), 500 mM NaCl, 10% glycerol,1 mM DTT, and Pierce Protease Inhibitor Tablets (Fisher Scientific). Cells were lysed with sonication, and the cleared lysate was loaded onto Strep-Tactin Superflow Plus (Qiagen). After extensive washing with the lysis buffer, proteins were eluted with lysis buffer supplemented with 2.5 mM desthiobiotin (Sigma). The eluted SUMO-tagged proteins were digested by homemade Ulp1-R3 SUMO protease (Addgene: #113671)[48] at 1:100 ratio at 4 °C for 16 h. The SUMO protease and undigested proteins were removed by loading onto Ni-NTA resin. Target proteins unbound to the resin were concentrated and loaded onto a Superdex 200 Increase column (Cytiva) with buffer of 50 mM Tris-HCl (pH 7.5), 500 mM NaCl, 5% glycerol and 2 mM DTT. Fractions were concentrated and stored at –80 °C.

To prepare the TnsC–TniQ complex, two plasmids encoding TniQ and TnsC, respectively, were co-transformed to BL21-CodonPlus (DE3)-RIPL competent cells. Protein expression and purification procedures were the same as those for purifying TnsC.

In vitro transposition assay

Donor plasmid (pDonor(PmcCAST)) and target plasmid (pTarget(PmcCAST)) were gifts from Feng Zhang (Addgene #168162 and #168163, respectively). TnsAB, the TnsC–TniQ complex and *Pmc*Cascade were diluted to 5 μM with 25 mM Tris-HCl, pH 8.0, 500 mM NaCl, 1 mM EDTA, 1 mM DTT, and 25% glycerol. 200 nM of each protein, 100 ng pTarget, and 100 ng pDonor were added to the reaction buffer containing 26 mM HEPES (pH 7.5), 4.2 mM Tris-HCl pH 8.0, 2.1 mM DTT, 0.05 mM EDTA, 0.2 mM $MgCl_2$, 28 mM NaCl, 21 mM KCl, 1.35% glycerol, 50 μg/mL BSA, and 2 mM ATP (final pH 7.5) for a total volume of 20 μL based on an *in vitro* Tn7 transposition protocol.[9] Reactions were incubated at 30 °C for 30 min before being supplemented with 25 mM MgOAc2 and incubated at 30 °C for another 2 hours. 1 μL of the final products was used for direct PCR readout.

To test the effects of Cascade mutants and spacer length on the transposition activity, 16 nM of Cascade variants was added to the reaction for qPCR. To test the effects of TniQ mutants on the transposition activity, 20 nM of TnsC–TniQ variants was added to the reaction for qPCR.

To test whether *Pmc*CAST results in simple insertions or cointegrates and the ratio between the two if it is a mixture, a donor plasmid with kanamycin resistance gene as the cargo of the transposon will be used for transposition assay.[4,11] Successful transposition will confer the target plasmid resistance to kanamycin. The donor plasmid contains R6K-γ origin, which requires protein pi (encoded by the gene pir) to initiate replication. The Top10 competent cell used for transformation after *in vitro* transposition assay does not have protein pi, so pDonor_KanR cannot replicate in this host. This allows us to obtain the transposition products with kanamycin selection for sequencing to confirm the exact insertion site and restriction enzyme digestion to test whether the resultant transposition product is simple insertion or cointegrate. Simple insertion and cointegrate result in different pattern of products after BamHI digestion that can be visualized by agarose gel. Similar methods were used to check the activity of mutants with mutant forms of proteins as listed replacing their wild-type components.

To test the role of ATP hydrolysis for transposition activity, *in vitro* transposition assays were performed by replacing ATP with AMP-PNP or replacing TnsC with $TnsC^{E172Q}$.

Mass spectrometry

Putative Cas11 protein from the gel band was reduced using 10 mM dithiothreitol, alkylated using 55 mM iodoacetamide (Sigma; I1149), and digested with 1 μg of trypsin (Thermo Scientific; 90058). Digested peptides were extracted by sonication in 5% trifluoroacetic acid/60% acetonitrile and filtered using Cellulose Acetate Spin-X columns (Corning; 8161). Peptides were dried and resuspended in 3% acetonitrile/0.1% formic acid for subsequent mass spectrometry analysis. Peptides were analyzed by Liquid Chromatography with tandem mass spectrometry (LC-MS/MS) using a Q Exactive HF Hybrid Quadrupole-Orbitrap mass spectrometer coupled with a Dionex Ultimate 3000 RSPC Nano system (Thermo Fisher Scientific) as described previously.[60] The LC-MS/MS data were analyzed using MaxQuant v2.0.3.0 software.[57] MaxQuant searches were performed against the Uniprot *E. coli* protein database supplemented with the protein sequences of *Peltigera membranacea cyanobiont 210A* Cas5–8. Searches were performed using either specific trypsin/P digestion with 4 max missed cleavages allowed or semi-specific trypsin/P digestion with free N-terminus; variable modifications were oxidation of methionine and acetylation of N-terminus; fixed modification was carbamidomethyl(Cys). Mostly default search settings were used, including main search peptide tolerance of 4.5 ppm and protein and peptide false discovery rate of 1%. Full search parameter settings can be found in Table S1.

E. coli transposition assay

150 ng of pTarget(PmcCAST), 215 ng of pDonor(PmcCAST) and 300 ng of pHelper(PmcCAST)_*Pmc*PSP1_ΔTnsD or the mutated pHelper(PmcCAST)_*Pmc*PSP1_ΔTnsD were transformed into BL21(DE3) cells using the Gene Pulser Xcell electroporation system (BIO-RAD). The cells were allowed to recover for 1.5 hours and plated on LB agar plate containing 100 μg/mL ampicillin, 50 μg/mL chloramphenicol and 50 μg/mL spectinomycin. A portion of the cells were scraped after 16-hour incubation at 37 °C and re-plated on the LB agar plate supplemented with 0.1 mM IPTG, ampicillin, chloramphenicol and spectinomycin. The plate was incubated for another 16 hours at 37 °C before the cells were subjected to mini-prep. The extracted plasmids were diluted to ~ 0.7 ng/μL before being analyzed by qPCR. pHelper(PmcCAST)_entry_ΔTnsD which does not express crRNA were used as a negative control.

Polymerase Chain Reaction (PCR)

Forward primer pTarget_F, reverse primer pDonor_R (Table S5), and *in vitro* transposition reaction product were mixed to a final volume of 25 μL for PCR reactions with Phusion Hot Start II High-Fidelity PCR Master Mix (Thermo Scientific, # F565L). Cycling conditions were as follows: 1 cycle, 98°C, 3 min; 35 cycles, 98°C, 10 s, 66°C, 30 s, 72°C, 10 s; 1 cycle, 72°C, 1 min.

Quantitative PCR

Quantitative PCR reactions were performed using TaqMan Fast Advanced Master Mix (Applied Biosystems) using Bio-Rad CFX Connect Real-Time PCR Detection system. To test the amount of inserted plasmid in products after *in vitro* transposition, 2 μL of the product, 900 nM forward primer, 900 nM reverse primer, and 250 nM of LE probe were mixed. Thermal cycling conditions were as follows: 1 cycle, 50 °C, 2 min; 1 cycle, 95 °C, 1.5 min; 39 cycles, 95 °C, 10 s, 56 °C, 30 s; End. To test the amount of target plasmid in products after *in vitro* transposition, the samples were diluted by 100 times. 2 μL of the diluted samples, 900 nM forward primer, 900 nM reverse primer, and 250 nM of target plasmid probe were mixed. Thermal cycling conditions were as follows: 1 cycle, 50 °C, 2 min; 1 cycle, 95 °C, 1.5 min; 39 cycles, 95 °C, 10 s, 59 °C, 30 s; End. The insertion frequency was calculated as the copy number of inserted plasmid divided by the copy number of target plasmid. The sequences of all the primers and probes used are listed in Table S5.

Electrophoretic mobility shift assay (EMSA)

Electrophoretic mobility shift assays were carried out using purified Cascade and its mutants. The concentrations of Cascade and its mutants were determined using Bradford protein assay. To test the DNA binding affinity, 20 nM 5'-FAM labeled substrate DNA (Table S5) was incubated with increasing concentrations (10, 20, 40, 60, 80, 100, 200, 400, 600, 800, 1000 nM) of purified Cascade or mutants at 37 °C for 30 min in buffer containing 20 mM Tris-HCl (pH 7.5), 200 mM NaCl, 5 mM $MgCl_2$ and 5% glycerol. 10 μL of binding reaction product were resolved on a native 5% TBE polyacrylamide gel at room temperature under 100 V for 45–55 minutes. The gels were visualized by fluorescence imaging (GE Amersham Imager 600 system). Quantification of EMSA data was performed by measuring the intensities of unbound DNA bands using ImageJ,[56] with DNA input calibrated to 100%. To determine binding affinity, binding curves were fitted using the 'specific binding with Hill slope' function in GraphPad Prism.

Preparation of the *Pmc*Cascade–DNA complex

The TS and NTS DNA oligos (Table S5) were synthesized by IDT and annealed at 1:1 molar ratio. *Pmc*Cascade was incubated with target DNA or PAM only DNA at a ratio of 1:1.8 at 37 °C for 30 min in buffer containing 20 mM HEPES (pH 7.5), 200 mM NaCl, 1 mM DTT, and 5 mM $MgCl_2$. After incubation, the sample was loaded onto a Superdex 200 Increase column (Cytiva). Fractions were analyzed by SDS-PAGE and 12% TBE-Urea gel (Novex) before concentrated to 0.5 mg/mL.

Preparation of the TnsC oligomer

400 μL of TnsC at 3 mg/mL in buffer containing 50 mM Tris-HCl (pH 7.5), 500 mM NaCl, 5% glycerol and 1 mM DTT was added to 500 μL of buffer containing 20 mM HEPES (pH 7.5), 200 mM NaCl, 5 mM $MgCl_2$, 2% glycerol, 1 mM DTT, and 2 mM ATP. Sample was incubated on ice for 30 min before being loaded to a Superdex 200 Increase column (Cytiva) equilibrated with buffer containing 20 mM HEPES (pH 7.5), 200 mM NaCl, 5 mM $MgCl_2$, 2% glycerol, 1 mM DTT, and 0.3 mM ATP. The peak fractions were analyzed by SDS-PAGE and concentrated to 2 mg/mL. To examine the oligomerization state of TnsC in the presence of ADP, a similar experiment was conducted with the exception of replacing 2 mM ATP with 2 mM ADP.

Preparation of TnsC–DNA and TnsC–TniQ–DNA

To assemble the TnsC–DNA complex, TnsC with a His-Strep-SUMO tag was purified with Strep-Tactin Superflow Plus (Qiagen) column in high-salt buffer containing 50 mM Tris-HCl pH 8.0, 1 M NaCl, 10% glycerol, 5 mM EDTA, 1 mM DTT. 100 μL 60-bp DNA at 45 μM (annealed using TnsC_DNA_60_1 and TnsC_DNA_60_2, Table S5) was added to 4 mL TnsC at 23 μM at a ratio of 1:20 (DNA:TnsC). Homemade Ulp1-R3 SUMO protease was also added at a ratio of 1:100 (protease:TnsC) to remove the tag on TnsC. The His-Strep-SUMO-TnsC, DNA and SUMO protease mixture was dialyzed against buffer containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 2% glycerol, 5 mM $MgCl_2$, 1 mM ATP at 4 °C overnight. The sample was filtered to remove precipitate, concentrated to 1 mL, and loaded to a Superdex 200 Increase column (Cytiva) equilibrated with buffer containing 20 mM HEPES (pH 7.5), 200 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, and 0.3 mM ATP. The peak fractions were analyzed by SDS-PAGE and concentrated to 2 mg/mL.

The TnsC–TniQ–DNA complex was assembled in a similar fashion. The TnsC–TniQ complex was purified with Strep Tactin Superflow Plus (Qiagen) in 50 mM Tris-HCl pH 8.0, 1 M NaCl, 10% glycerol, 1 mM DTT. 67 μL 60-bp DNA at 45 μM was added to 3 mL TnsC–TniQ at 20 μM. Homemade Ulp1-R3 SUMO protease was added at a ratio of 1:100 (protease:TnsC) to remove the tag. The His-Strep-SUMO-TnsC-TniQ, DNA and SUMO protease mixture was dialyzed against buffer containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 2% glycerol, 5 mM $MgCl_2$, 1 mM ATP at 4 °C overnight. The sample was filtered to remove precipitate and concentrated to 1 mL before loaded to a Superdex 200 Increase column (Cytiva) equilibrated with buffer containing 20 mM HEPES (pH 7.5), 200 mM NaCl, 5 mM $MgCl_2$, 1 mM DTT, and 0.3 mM ATP. The peak fractions were analyzed by SDS-PAGE and concentrated to 2 mg/mL.

Preparation of *Pmc*Cascade–TnsC–TniQ–DNA

*Pmc*Cascade used to assemble *Pmc*Cascade–TnsC–TniQ–DNA complex was purified with Ni-NTA resin (Qiagen) followed by HiTrap Q HP (Cytiva) as described above. The TnsC–TniQ complex was purified with Strep Tactin Superflow Plus (Qiagen) resin followed by SUMO protease cleavage and reverse-IMAC using Ni-NTA resin (Qiagen). After adjusting the buffer to 50 mM Tris-HCl (pH7.5), 150 mM NaCl, 2% glycerol, 5 mM $MgCl_2$ and 1 mM ATP, 700 μL *Pmc*Cascade at 0.9 μM was incubated with 35 μL 85-bp target DNA substrate at 45 μM (annealed by using NTS 85 nt and TS 85 nt; Table S5) at 37 °C for 30 min to assemble the *Pmc*Cascade–DNA complex. The resulting mixture was added to 30 mL of 1.6 μM TnsC–TniQ complex in buffer containing 50 mM Tris-HCl (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, 2% glycerol, and 1 mM ATP, and incubated at 4 °C overnight. The sample was later loaded onto 1 mL Ni-NTA resin for four times, washed by buffer containing 20 mM HEPES (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, 0.3 mM ATP, and 20 mM imidazole and eluted by buffer containing 20 mM HEPES (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, 0.3 mM ATP and 200 mM imidazole. The eluted sample was analyzed by SDS-PAGE and concentrated to 0.5 mg/mL. To investigate the necessity of target DNA in the recruitment of TnsC and TniQ, a similar pull-down assay was performed without adding target DNA substrate. Non-specific binding between TnsC and Ni-NTA resin was investigated under the same experimental conditions, using BSA as a control.

Electron microscopy

Aliquots of 3 μL *Pmc*Cascade–TnsC–TniQ–DNA sample (or *Pmc*Cascade, or *Pmc*Cascade–PAM, or *Pmc*Cascade–DNA) at ~ 0.5 mg/mL were applied to glow-discharged UltrAuFoil holey gold grids (R1.2/1.3, 300 mesh). The grids were blotted for 2.5 seconds and plunged into liquid ethane using a ThermoFisher Scientific Mark IV Vitrobot. Cryo-EM data were collected with a Titan Krios G4 microscope (FEI) operated at 300 kV and images were collected using EPU at a nominal magnification of 81,000x (resulting in a calibrated physical pixel size of 1.05 Å/pixel) with a defocus range of 0.8–2.0 μm. The images were recorded on a K3 electron direct detector in super-resolution mode at the end of a GIF-Quantum energy filter operated with a slit width of 20 eV. A dose rate of 20 electrons per pixel per second and an exposure time of 3.12 seconds were used, generating 40 movie frames with a total dose of ~ 54 electrons per $Å^2$. Statistics for cryo-EM data are listed in Table S3.

TnsC, TnsC–DNA and TnsC–TniQ–DNA were studied similarly, except for that 2 mg/mL samples were used for grid preparation.

Image processing

For the *Pmc*Cascade–TnsC–TniQ–DNA dataset, movie frames were imported to cryoSPARC.[53] Movie frames were aligned by patch motion correction with a binning factor of 2. Contrast transfer function (CTF) parameters were estimated using Patch CTF. A few thousand particles were auto-picked without a template to generate 2D averages for subsequent template-based auto-picking. The auto-picked and extracted particles were screened by 2D classification to exclude false and bad particles that fall into 2D averages with poor features. A subset of the particles (100,000 particles) were used to generate initial models in cryoSPARC. All screened particles were subjected to heterogenous refinement using the initial models. Particles in each 3D class were subject to homogenous

refinement. The resolution was estimated based on the Fourier Shell Correlation (FSC) = 0.143 criterion. Finally, focused refinements were further performed to improve local resolutions. A combined map was calculated by the *vop maximum* program in Chimera[49] and used for model building. To make *Pmc*Cascade–DNA–TniQ–TnsC composite, maps of *Pmc*Cascade–DNA–TniQ–TnsC and TnsC–TniQ–DNA were superimposed according to the densities of TniQ, and a composite map was generated by the *vop maximum* program in Chimera.[49] A workflow for data processing is summarized in Figure S2. A similar workflow was used for *Pmc*Cascade, *Pmc*Cascade–PAM, *Pmc*Cascade–DNA, TnsC, TnsC–DNA and TnsC–TniQ–DNA.

Model building and refinement

Structures of all components were first predicted using AlphaFold. The individual structures were fitted into the maps as a rigid-body in UCSF Chimera[49] and manually adjusted in COOT.[52] Structure of the RNA hairpin of the crRNA was predicted by RNAComposer[58] and fitted into the maps as a rigid-body in UCSF Chimera.[49] The remaining part of the crRNA and target DNA were built *de novo* in COOT. The model was then fitted in COOT with all molecule self-restrains to maintain hydrogen bonds and stacking interactions between bases. Finally, refinement of the structure models against corresponding maps were performed using *phenix.real_space_refine* tool in Phenix.[51]

Sequence alignment

Protein sequence alignment shown in Figure S6C was performed by Clustal Omega.[59] The alignment diagram was plotted using ESPript.[54] Sequence identities and similarities for Table S4 were calculated the Sequence Manipulation Suite.[55]

Structural visualization

Figures were generated using PyMOL, UCSF Chimera[49] and ChimeraX.[50]

QUANTIFICATION AND STATISTICAL ANALYSIS

Data were processed and plotted using GraphPad Prism 8. Statistical validation for the final structural models shown in Table S3 was performed using Phenix.[51]

Supplemental figures

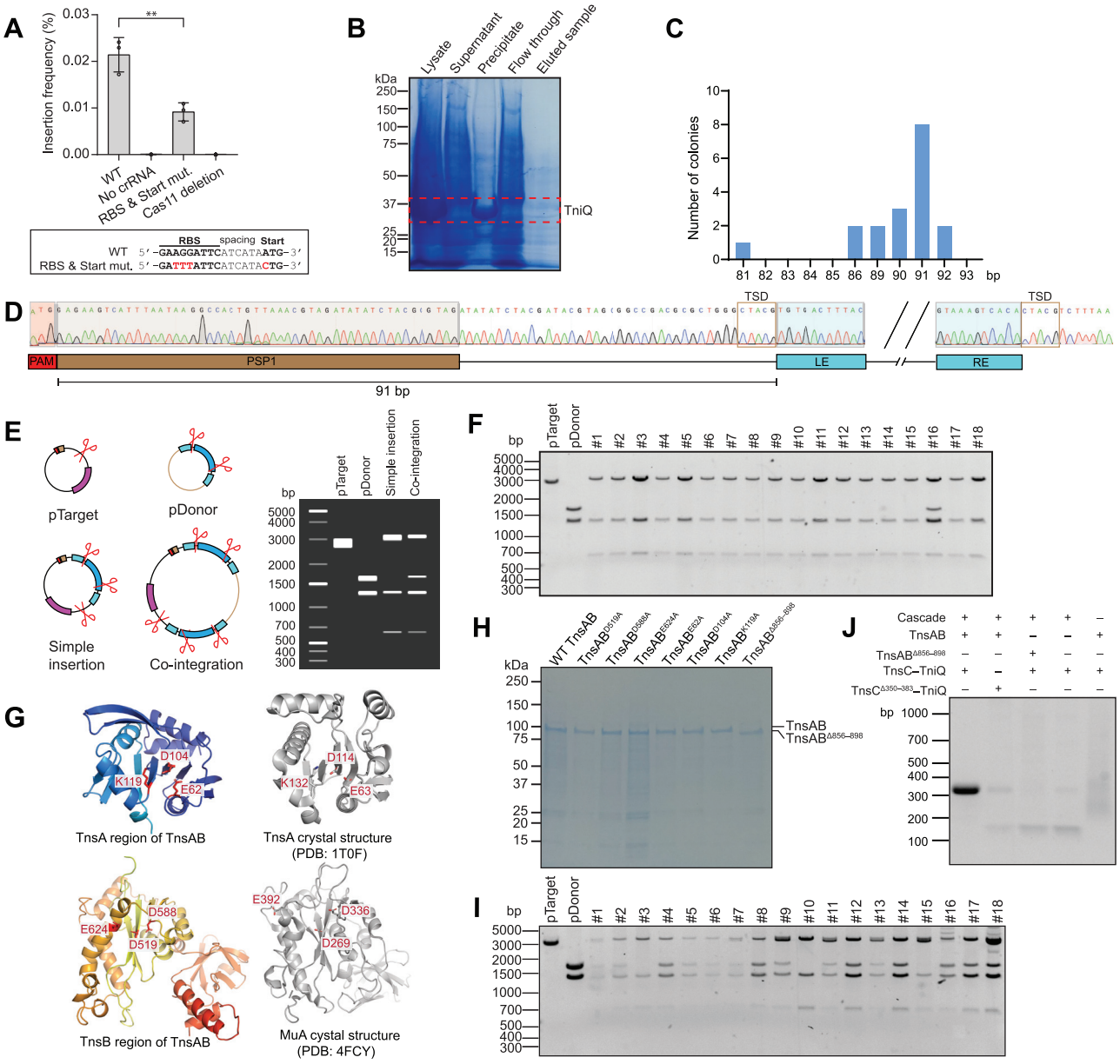


Figure S1. Biochemical reconstitution of RNA-guided DNA transposition in *Pmc*CAST, related to Figure 1

(A) *E. coli* transposition assay using wild-type and mutant Cas11. When the RBS and the start codon of Cas11 are mutated, some transposition activity remains. Possible sources of this remaining activity could be an incomplete disruption of Cas11 translation, the existence of alternative translation sites for Cas11 (Table S2), and a potential compensatory functionality of Cas8 through its C-terminal region. Data are presented as mean values ± SD (n = 3). Two-sided t test was performed (**$p < 0.01$).

(B) SDS-PAGE gel for the purification of TniQ alone.

(C) Distribution of insertion positions relative to the PAM sequence.

(D) One example of the sequencing result showing the insertion site at a position 91-bp after the PAM sequence.

(E) Simulated results of BamHI digestion of pTarget, pDonor, simple insertion, and cointegration products.

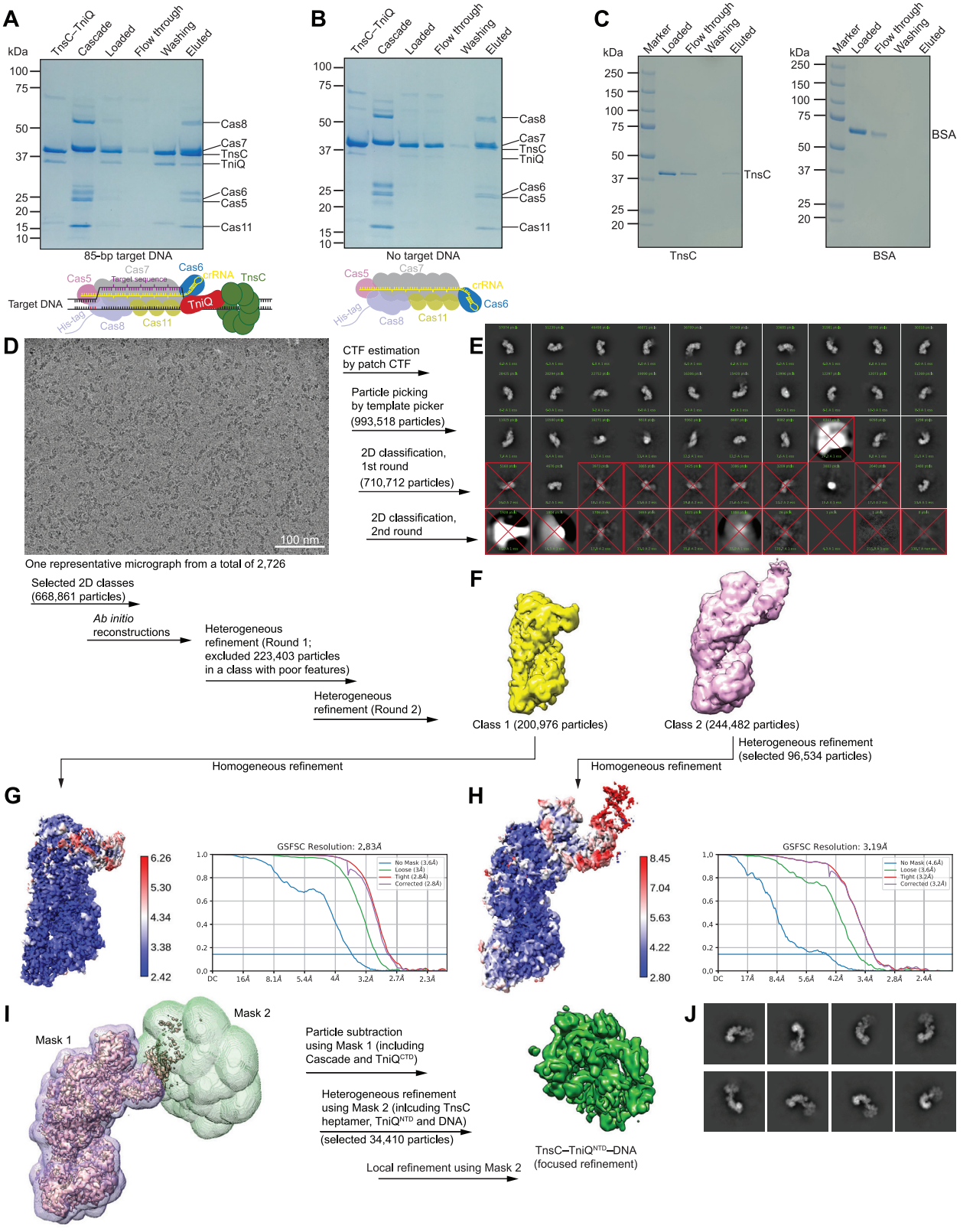
(F) BamHI digestion of plasmids from 18 randomly picked colonies using wild-type TnsAB.

(G) Close-up view of the TnsA and TnsB regions of the TnsAB AlphaFold model with catalytic residues colored in red. Shown on the right are the TnsA crystal structure from Tn7 and the crystal structure of MuA for comparison.

(H) SDS-PAGE of purified wild-type and mutant TnsAB.

(I) BamHI digestion of plasmids extracted from 18 randomly picked colonies using TnsAB with TnsA inactivated.

(J) PCR readout of *in vitro* transposition assay with C-terminal deletions from TnsC ($TnsC^{\Delta 350–383}$) and TnsAB ($TnsAB^{\Delta 856–898}$).



(legend on next page)

Figure S2. Assembly and cryo-EM data processing of *Pmc*Cascade-DNA-TniQ-TnsC, related to Figure 2

(A and B) SDS-PAGE gel for the assembly of *Pmc*Cascade-DNA-TniQ-TnsC. Shown in (B) is a control where the target DNA is omitted. Cartoon diagrams of the expected eluted complexes are shown at the bottom.

(C) Examination of nonspecific binding between TnsC and Ni-NTA beads (left) and between BSA and Ni-NTA beads (right).

(D) A representative raw cryo-EM micrograph of *Pmc*Cascade-DNA-TniQ-TnsC from a total of 2,726 micrographs.

(E) 2D class averages of selected particles. Crossed classes were excluded from further analysis.

(F) Two 3D reconstructions from heterogeneous refinement.

(G) Local resolution map of *Pmc*Cascade-DNA^{P23}. Global gold-standard FSC is shown on the right.

(H) Local resolution map of *Pmc*Cascade-DNA-TniQ-TnsC. Global gold-standard FSC is shown on the right.

(I) Strategies for improving the resolution of the TniQ^{NTD}-TnsC-DNA region in *Pmc*Cascade-DNA-TniQ-TnsC by particle subtraction and local refinement.

(J) Representative 2D class averages of selected particles (prior to signal subtraction) for the reconstruction of the TniQ^{NTD}-TnsC-DNA region as indicated in (I).

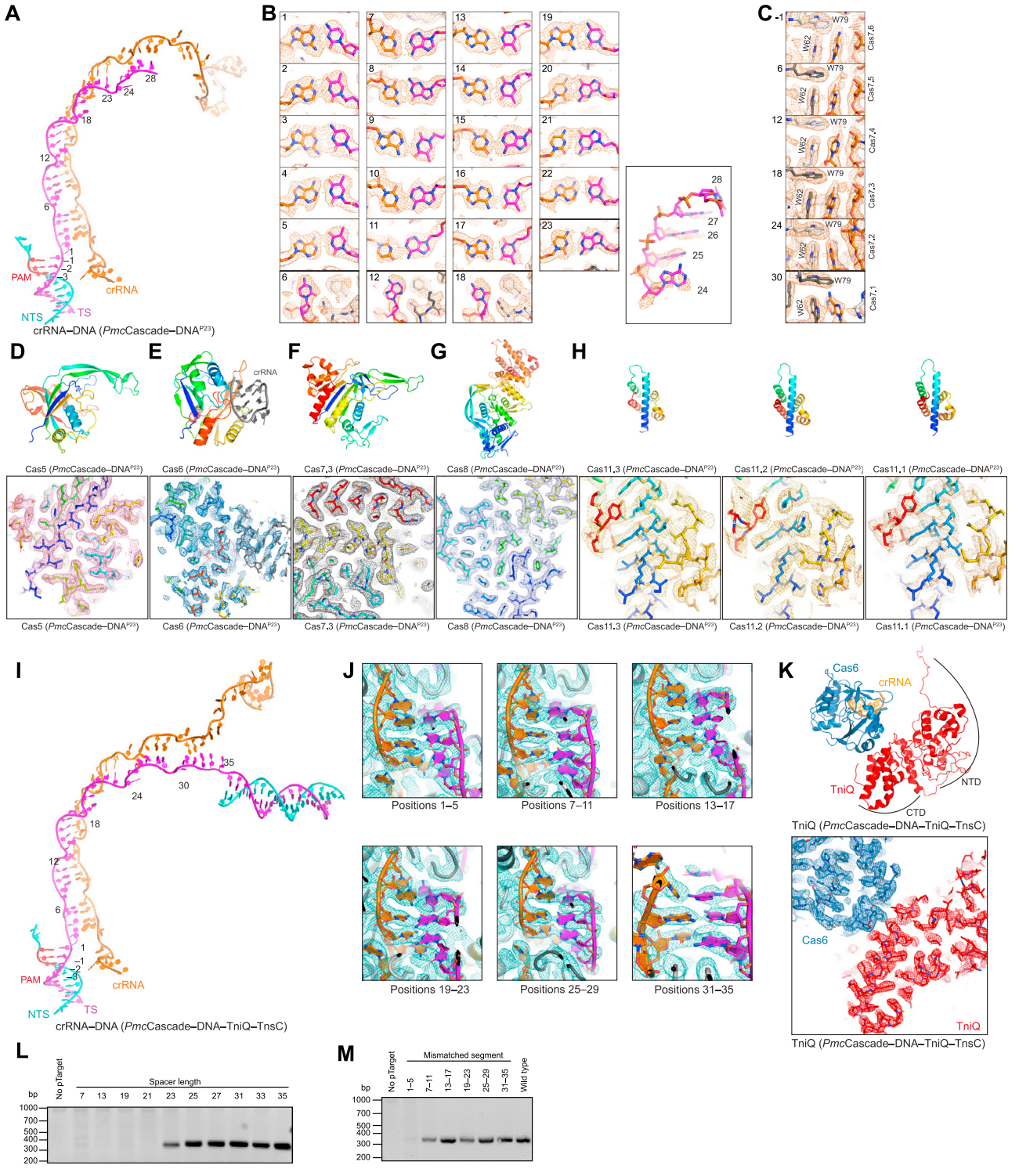


Figure S3. Detailed cryo-EM density map of *Pmc*Cascade-DNA^{P23} and *Pmc*Cascade-DNA-TniQ-TnsC, related to Figures 3 and 4

(A) Atomic model of crRNA-DNA in *Pmc*Cascade-DNA^{P23}.

(B) Fitting between the crRNA-DNA model and the cryo-EM map of *Pmc*Cascade-DNA^{P23}.

(C) Fitting between the crRNA kink nucleotides and the cryo-EM map of *Pmc*Cascade-DNA^{P23}.

(D–H) Structural model of Cascade subunits and their fitting to the cryo-EM map of *Pmc*Cascade-DNA^{P23}.

(I) Atomic model of crRNA-DNA in *Pmc*Cascade-DNA-TniQ-TnsC.

(legend continued on next page)

(J) Fitting between the crRNA-DNA model and the cryo-EM map of *Pmc*Cascade-DNA-TniQ-TnsC.
(K) Structural model of the TniQ^{CTD}-Cas6-crRNA hairpin and its fitting to the cryo-EM map of *Pmc*Cascade-DNA-TniQ-TnsC.
(L) PCR readout of *in vitro* transposition assay using target plasmids with different spacer lengths.
(M) PCR readout of *in vitro* transposition assay using target plasmids with mismatches at each 5-nt segment.

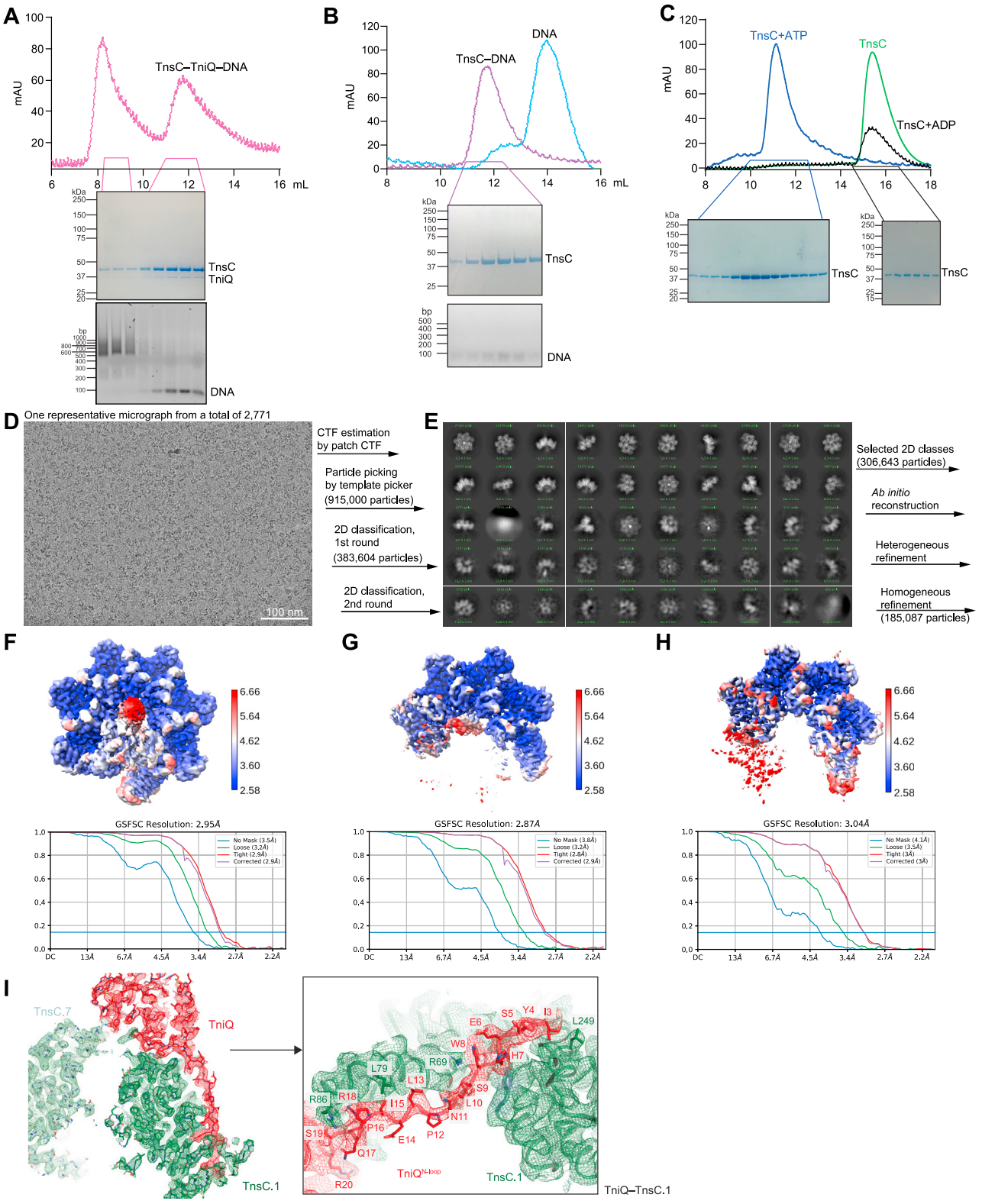


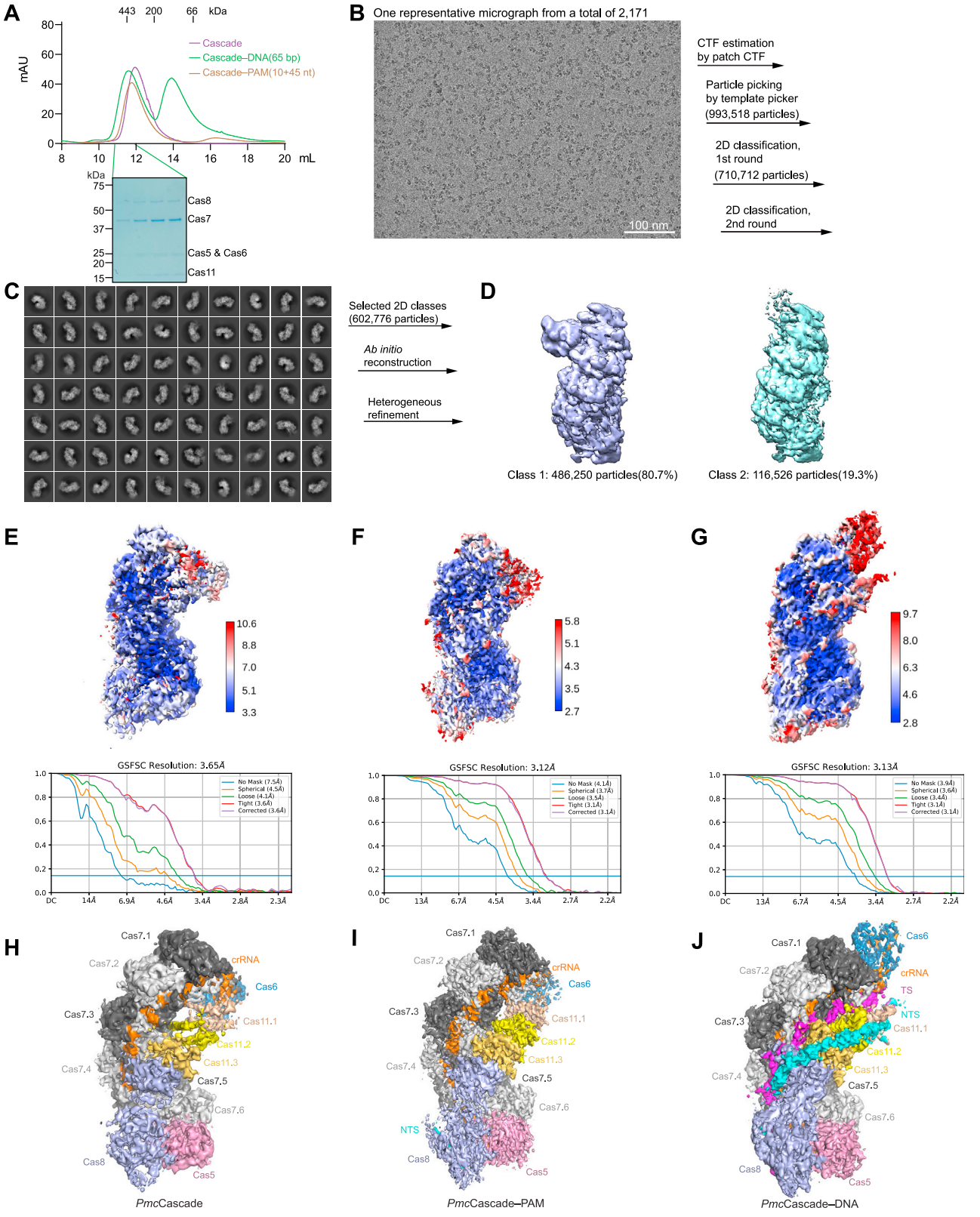
Figure S4. Cryo-EM analysis for TnsC-TniQ-DNA, TnsC-DNA, and TnsC, related to Figures 2 and 5

(A) SEC profile for the purification of TnsC-TniQ-DNA. Shown below are SDS-PAGE and agarose gels of the indicated fractions.

(B) SEC profiles for the purification of TnsC-DNA. SDS-PAGE and agarose gels of the indicated fractions are shown.

(legend continued on next page)

(C) SEC profiles for TnsC alone, TnsC with ATP, and TnsC with ADP. SDS-PAGE gels of the indicated fractions are shown.
(D) A representative raw cryo-EM micrograph of TnsC-TniQ-DNA from a total of 2,771 micrographs.
(E) 2D class averages of selected particles for TnsC-TniQ-DNA.
(F) Local resolution map of TnsC-TniQ-DNA. Global gold-standard FSC is shown at the bottom.
(G) Local resolution map of TnsC-DNA. Global gold-standard FSC is shown at the bottom.
(H) Local resolution map of TnsC in ATP-bound state. Global gold-standard FSC is shown at the bottom.
(I) Interactions between TniQ and TnsC in TnsC-TniQ-DNA. Cryo-EM map is shown in mesh.



(legend on next page)

Figure S5. Cryo-EM analysis for *Pmc*Cascade complexes, related to Figure 3

(A) SEC profiles for *Pmc*Cascade, *Pmc*Cascade-PAM, and *Pmc*Cascade-DNA.

(B) A representative raw cryo-EM micrograph of *Pmc*Cascade-DNA from a total of 2,171 micrographs.

(C) Representative 2D class averages from a total of 100 class averages.

(D) Two major classes from 3D classification. Class 1 represents the *Pmc*Cascade-DNA^{P23} state, whereas class 2 represents the *Pmc*Cascade-DNA state where full R-loop is formed.

(E) Local resolution map of *Pmc*Cascade. Global gold-standard FSC is shown at the bottom.

(F) Local resolution map of *Pmc*Cascade-PAM. Global gold-standard FSC is shown at the bottom.

(G) Local resolution map of *Pmc*Cascade-DNA. Global gold-standard FSC is shown at the bottom.

(H–J) Cryo-EM density maps of *Pmc*Cascade, *Pmc*Cascade-PAM, and *Pmc*Cascade-DNA (full R-loop) at 3.6, 3.1, and 3.1 Å, respectively.

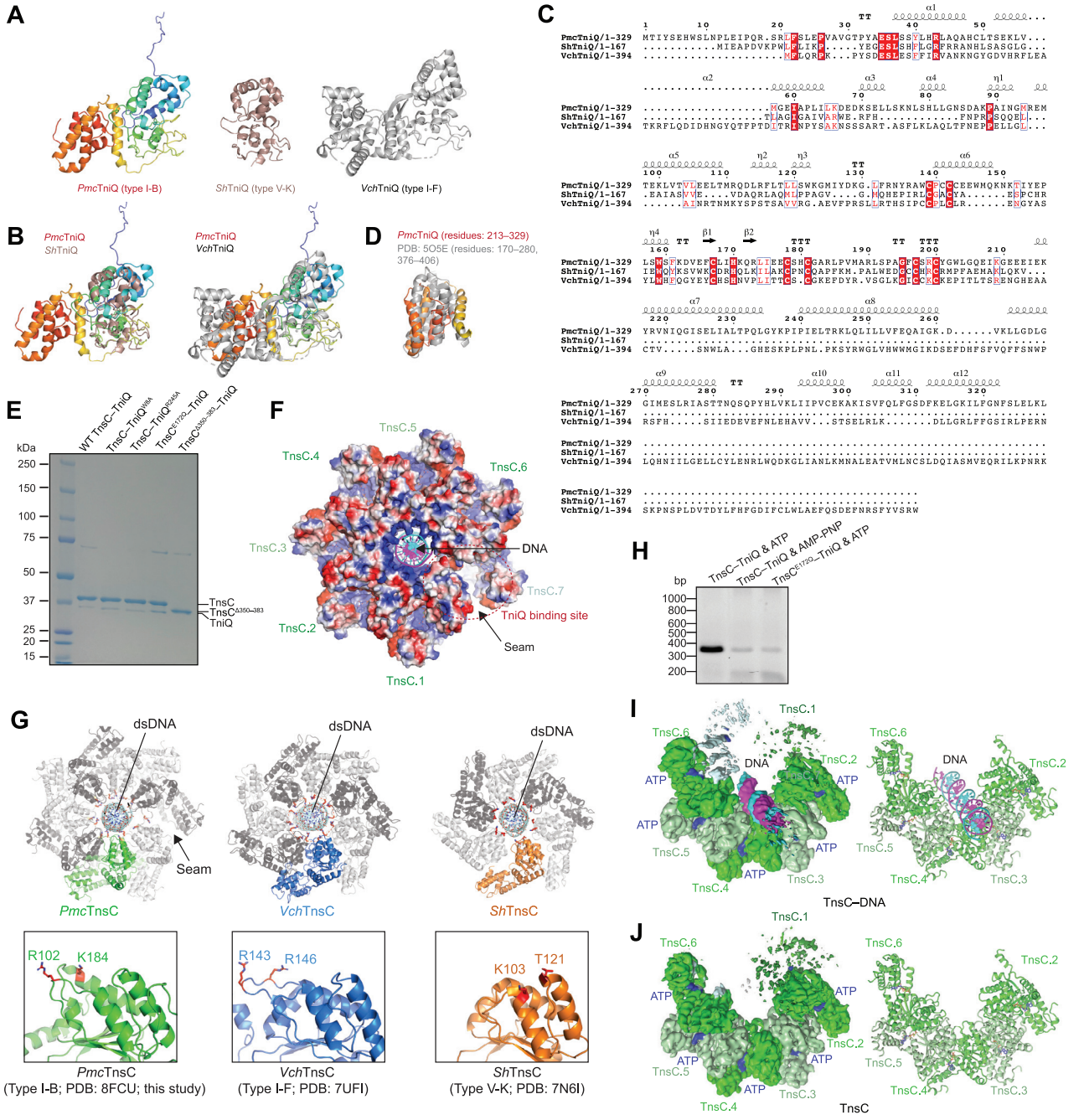


Figure S6. Structural analysis of TniQ and TnsC, related to Figure 6

(A) Structures of TniQ from *Pmc*CAST, *Sh*CAST, and *Vch*CAST.

(B) Structural superimposition between *Pmc*TniQ and *Sh*TniQ, and between *Sh*TniQ and *Vch*TniQ (only one TniQ from the *Vch*TniQ dimer is displayed for clarity).

(C) Sequence alignment of TniQ from *Pmc*CAST, *Sh*CAST, and *Vch*CAST systems.

(D) A DALI search revealed the closest match of the CTD domain of *Pmc*TniQ as the UDP-N-acetylglucosamine-dolichyl-phosphate N-acetylglucosaminephosphotransferase (DPAGT1) (PDB: 5O5E; *Z* score, 6.3), although there is no apparent functional linkage between DPAGT1 and TniQ.

(E) SDS-PAGE gel of purified wild-type and mutant TnsC-TniQ.

(F) DNA interacting with the central positively charged pore of the TnsC heptamer, with TnsC displayed as an electrostatic potential surface based on the TnsC-TniQ-DNA structure.

(legend continued on next page)

(G) Comparison of the TnsC-DNA structures from the *Pmc*CAST, *Vch*CAST, and *Sh*CAST systems. Top: side-by-side comparison of the TnsC-DNA structures, with key residues in the central pores highlighted in red. Bottom: close-up comparison of a single TnsC subunit extracted from the TnsC oligomers.
(H) PCR readout of *in vitro* transposition assay when ATP was substituted with AMP-PNP or TnsC was replaced with $TnsC^{E172Q}$.
(I) Cryo-EM map and atomic model of TnsC-DNA.
(J) Cryo-EM map and atomic model of TnsC in ATP-bound state.

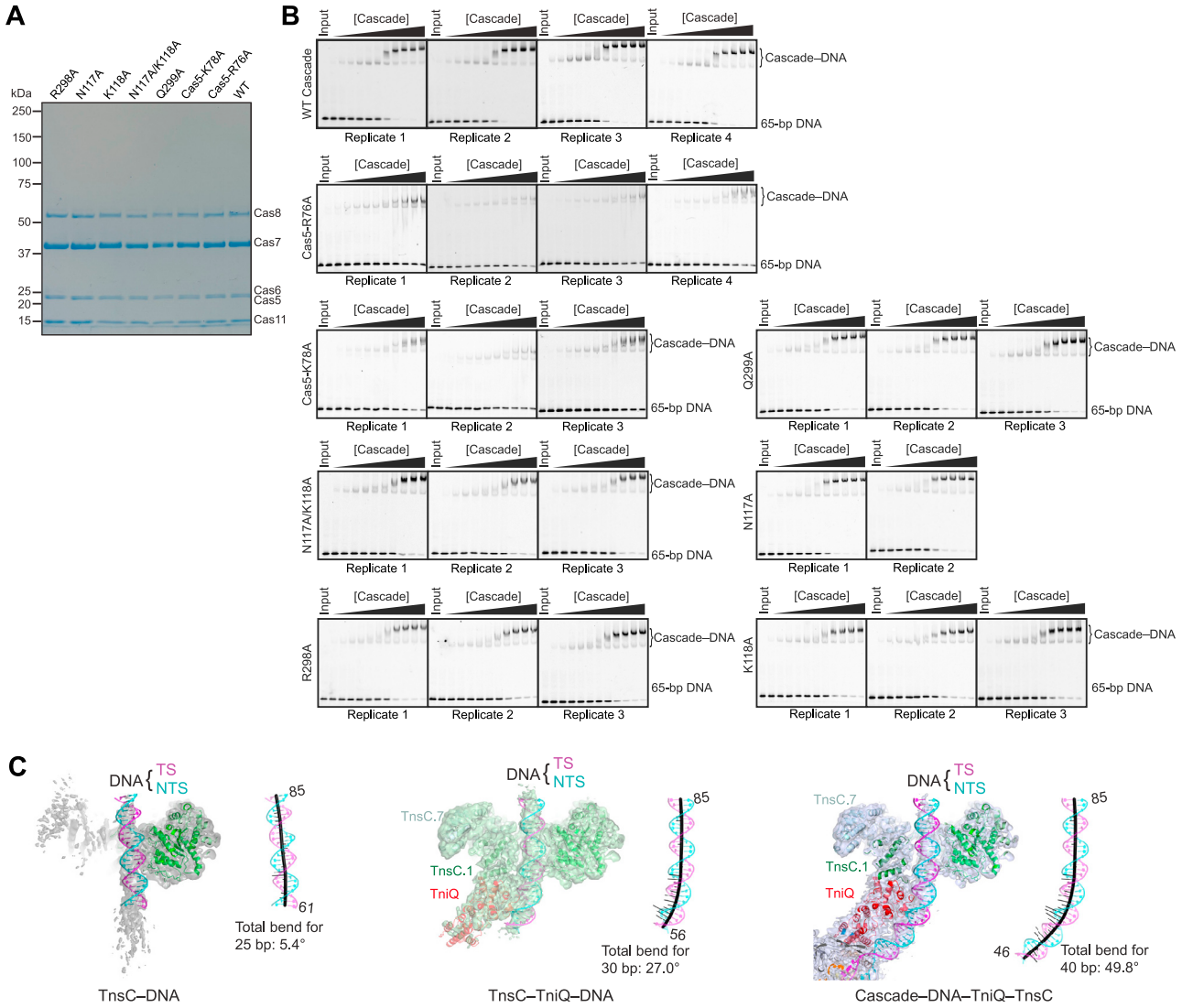


Figure S7. Target DNA recognition in *Pmc*CAST, related to Figure 7

(A) SDS-PAGE gel of purified wild-type and mutant Cascade.

(B) EMSA results showing the binding of wild-type and mutant Cascade to the target DNA.

(C) DNA distortions in TnsC-DNA, TnsC-TniQ-DNA, and *Pmc*Cascade-DNA-TniQ-TnsC. Shown on the left for each structure is the structural model with cryo-EM map in transparent surface. Shown on the right is the DNA model in each structure with the axis shown as a line (calculated using the Curves+ program).