Cell Research

www.nature.com/cr
www.cell-research.com

LETTER TO THE EDITOR

Cryo-EM structure of a type I-F CRISPR RNA guided surveillance complex bound to transposition protein TniQ

Cell Research (2020) 0:1–3; https://doi.org/10.1038/s41422-019-0268-y

Dear Editor,

CRISPR-Cas systems provide adaptive immunity in prokaryotes through RNA-guided cleavage of mobile genetic elements and have been harnessed as powerful genome editing tools. A programmable guide RNA and the effector Cas protein are delivered to mammalian cells to create double strand breaks on the target sequence in the genome, triggering endogenous DNA repair pathways that enable genetic changes. However, recent studies revealed transposon-associated CRISPR-Cas systems that are inactive in target cleavage but direct transposases for RNA-guided DNA transposition.[1–4] Those findings open a new paradigm for precise DNA insertion independent of DNA repair pathways.[1,2] RNA-guided DNA integration in type I-F CRISPR-Cas system from *Vibrio cholerae* requires the DNA-targeting Cascade complex (also called the Csy complex) that is composed of Cas6, Cas7, a natural fusion of Cas8 and Cas5 (Cas8/5), and a 60-nt crRNA, along with the transposition machinery including TnsA, TnsB, TnsC and TniQ (a homolog of TnsD) (Fig. 1a). TniQ binds to the *Vibrio cholerae* Cascade complex (VcCascade) and recruits the core TnsABC machinery for transposition, thus playing an essential role in transposition. To understand the molecular basis if the RNA-guided DNA integration process, we determined the structure of VcCascade bound to TniQ at 3.1 Å resolution by cryo-EM (Fig. 1b, c; Supplementary information, Figs. S1–3, and Table S1). The resultant maps are of sufficient quality for accurate atomic model building for VcCascade and TniQ, except for the helical bundle of Cas8/5 which is built as a poly-alanine model (Fig. 1b).

The architecture of VcCascade adopts a "G" shape (Fig. 1b), similar to that of Cascade from *Pseudomonas aeruginosa* (PaCascade)[5–8] (Supplementary Information Fig. S4a–d). Cas6, Cas7 and Cas8/5 with a stoichiometry of 1:6:1 are integrated by the crRNA, where the spacer sequence is flanked by a 3′ hairpin structure and a 5′ handle (Supplementary information, Fig. S4e). Six copies of Cas7 (Cas7.1–Cas7.6) assemble into a right-handed helical backbone with each Cas7 corresponding to a twist of ~46° and a rise of ~13 Å (Fig. 1b, c; Supplementary information, Fig. S4f). The structure of Cas7 is similar to that of Cas7 from PaCascade, with a root-mean-square deviation (RMSD) of 1.6 Å (Supplementary information, Fig. S4g). The Cas7 backbone is capped by Cas6 associated with the crRNA 3′ hairpin, forming the head of VcCascade (Fig. 1b, c; Supplementary information, Fig. S4h). Although the overall folds of Cas6 from VcCascade and PaCascade are similar, substantial differences were observed (RMSD = 5.6 Å; Supplementary information, Fig. S4i), indicative of distinct functional roles (to be discussed below). In the tail of VcCascade, Cas7 backbone is terminated by binding of Cas8/5 to the crRNA 5′ handle (Fig. 1b, c; Supplementary Information Fig. S4j). Cas8/5 consists of an N-terminal domain (Cas8/5^{NTD}), a middle helical bundle (Cas8/5^{HB}) and a C-terminal domain (Cas8/5^{CTD}), which are homologous to Cas8^{NTD}, Cas8^{HB} and Cas5 in PaCascade, respectively (Supplementary information, Fig. S4j–m). Cas8/5^{HB} shares a similar fold to Cas8^{HB} of PaCascade (RMSD = 4.8 Å), and connects the tail and the head of VcCascade (Supplementary information, Fig. S4m).

TniQ folds into a bean-shaped structure that is divided into N-terminal (aa 1–188) and C-terminal (aa 193–394) halves connected by a short linker (Fig. 1d). The N-terminal half contains a helix-turn-helix domain (HTH1, α1–3), followed by an antiparallel β-sheet (β1–3) and a CCCH type zinc finger domain (ZnF)[9] (Fig. 1d; Supplementary information, Fig. S5a–d). The ZnF domain contains a zinc ion chelated by three cysteines (C128, C131 and C150) and a histidine (H153) which are invariant in TnsD (Fig. 1d; Supplementary information, Fig. S5c, d). The functions of ZnF domain include protein folding and assembly, and recognition of nucleic acid,[10] therefore the ZnF domain of TniQ is likely involved in recruitment of TnsABC or PAM-distal DNA association during transposition. The C-terminal half is composed of a helical domain (HD, α5–10), and a second HTH domain (HTH2, α10–14) that resembles the HTH domain of DNA-binding protein BldC[11] (RMSD = 1.3 Å) (Fig. 1d; Supplementary information, Fig. S5e).

TniQ forms a head-to-tail antiparallel dimer with a buried surface area of ~800 Å^2 (Fig. 1e; Supplementary information, Fig. S1f). The HTH2 domain of TniQ.2 is bound to a groove formed by HTH1 and the β-sheet of TniQ.1, and vice versa (Fig. 1e; Supplementary information, Fig. S5f, g). Both TniQ.1 and TniQ.2 interact with the head of VcCascade (Fig. 1f; Supplementary information Fig. S6a, b). TniQ.1 solely interacts with the palm domain of Cas7.1 through HTH1-mediated electrostatic interactions (Supplementary information, Fig. S6c), whereas TniQ.2 forms extensive contacts with VcCascade involving Cas6, Cas7.1, the Cas7.2 thumb domain, and Cas8/5^{HB} (Fig. 1g). Helix α8 and the α8-α9 loop of HD domain from TniQ.2 interacts with the ferredoxin-like domain of Cas6. In particular, F261, H265, V267 and V268 from TniQ insert into a hydrophobic pocket of the ferredoxin-like domain of Cas6 (Fig. 1g). GST pull-down assay shows reduced binding affinity between VcCascade and TniQ when V267 and V268 are mutated to glycine (Supplementary information, Fig. S6d, e). The HTH2 domain forms electrostatic interactions with Cas7.1 and the Cas7.2 thumb domain (Supplementary information, Fig. S6f). The HTH2 domain also engages Cas8/5^{HB}, although specific side-chain interactions are not determined due to the limited resolution in this region (Fig. 1b; Supplementary information, Fig. S6b). Cas8^{HB} recruits Cas2/3 for target DNA cleavage in PaCascade.[8] However, the gene encoding Cas2/3 is absent in the transposon-associated CRISPR-Cas systems,[1–4] suggesting a distinct role of Cas8/5^{HB}. Our structure shows that Cas8/5^{HB} in VcCascade is involved in TniQ recruitment, indicating a functional adaption for DNA transposition.

Structural comparison with PaCascade revealed significant rearrangements in Cas6 and Cas8/5^{HB} in VcCascade-TniQ. First, the binding mode of Cas6 and Cas7.1 is fundamentally different. In PaCascade, the ferredoxin-like domain of Cas6 contacts Cas7.1 extensively, exposing the thumb domain of Cas6 to solvent (Fig. 1h). In VcCascade, the association between Cas6 and Cas7.1 is

Received: 25 September 2019 Accepted: 11 December 2019
Published online: 03 January 2020

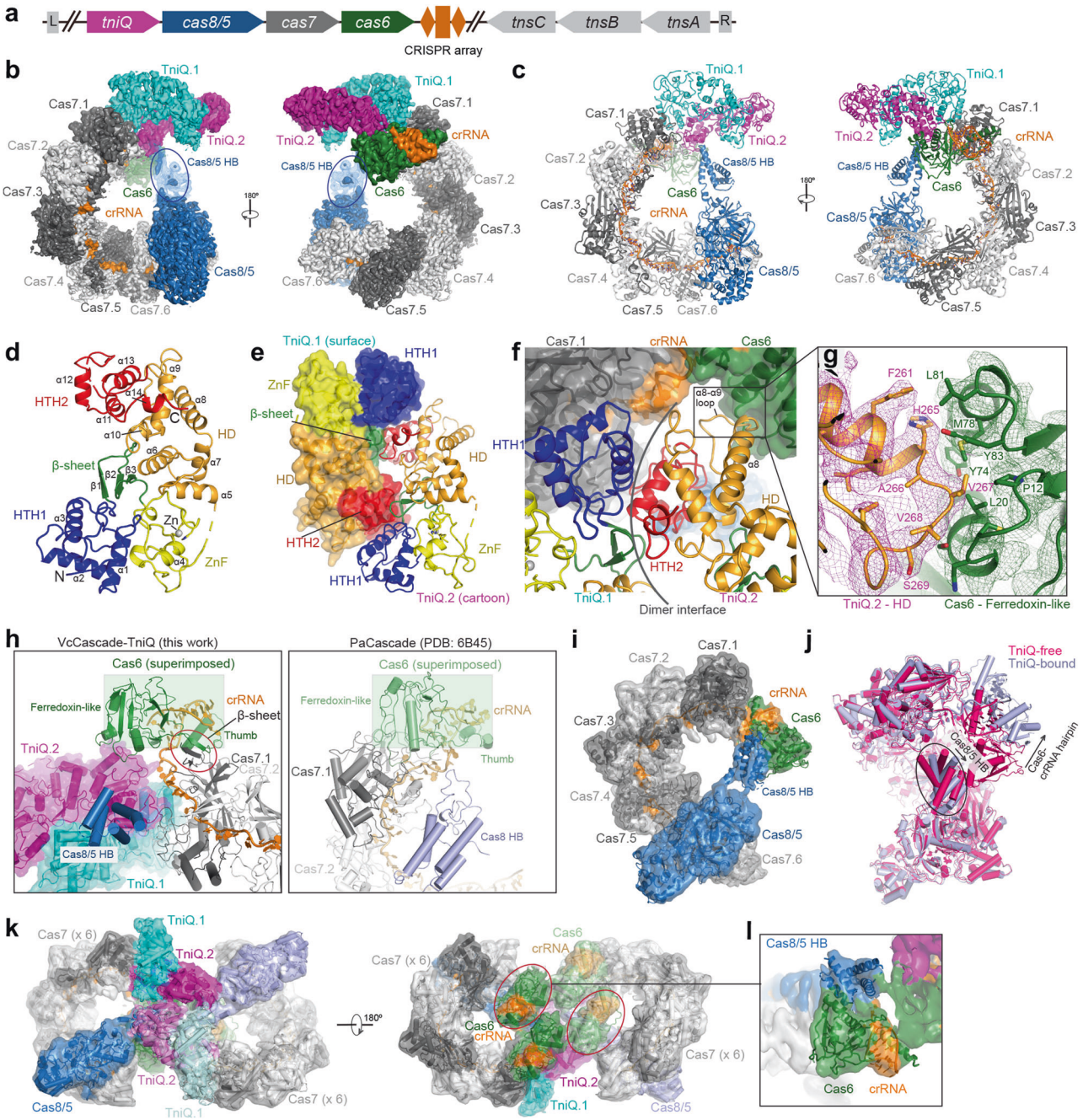

Fig. 1 Cryo-EM structure of VcCascade-TniQ. a Schematic diagram of the genomic organization of the Tn-7 transposon in *V. cholerae*. **b** Two views of the cryo-EM map of VcCascade-TniQ with each subunit color-coded. The density for Cas8/5^{HB} at lower threshold is shown in mesh. **c** Two views of the atomic model of VcCascade-TniQ in cartoon presentation. **d** Rainbow presentation of a TniQ subunit with each domain color-coded. **e** TniQ dimer shown in cartoon (TniQ.2) and surface (TniQ.1). **f** TniQ dimer binding site in VcCascade. The interface between two TniQ subunits is indicated. **g** Detailed interactions between the HD of TniQ.2 and the ferredoxin-like domain of Cas6. **h** Side-by-side comparison of VcCascade and PaCascade with Cas6 superimposed. In VcCascade, the thumb domain of Cas6 interacts with Cas7.1 involving the formation of a parallel intermolecular β-sheet (indicated by a red circle). **i** Cryo-EM map of TniQ-free VcCascade with structural model shown in cartoon. **j** Structural comparison between TniQ-free (hot pink) and TniQ-bound (light blue) VcCascade. Movements of Cas6-crRNA hairpin and Cas8/5^{HB} are indicated by arrows. **k** Top and bottom views of the cryo-EM map of the supercomplex with underlying model shown in cartoon. The supercomplex contains two copies of VcCascade-TniQ, and two additional copies of Cas6-crRNA 3′ hairpin (indicated by red circles). **l** Rigid-body fitting of Cas6-crRNA hairpin and Cas8/5 HB into the additional densities of the supercomplex.

mediated by the Cas6 thumb domain, making the ferredoxin-like domain of Cas6 accessible to TniQ (Fig. 1h). In addition, Cas8^{HB} is located in the center of PaCascade, contacting Cas7 backbone (Supplementary information, Fig. S4c). However, in VcCascade, Cas8/5^{HB} is displaced from the center and associated with the head (Supplementary information, Fig. S4a). Both Cas6 and Cas8/5^{HB} are involved in TniQ recruitment, thus the structural differences and specific interactions between VcCascade and TniQ may explain why VcCascade could direct DNA transposition, but PaCascade failed even when TniQ is fused to it.[1]

By 3D classification of particle images, we observed a population representing VcCascade alone, accounting for ~16%

of total particles. We determined the structure of TniQ-free VcCascade to 3.8 Å resolution (Fig. 1i). Structural comparison between TniQ-free and TniQ-bound VcCascade reveals that the Cas7 backbone and Cas8/5 (except for Cas8/5^{HB}) are almost the same. However, TniQ association induces a ~22 Å movement of Cas6 away from Cas8/5^{HB} accompanied by the formation of a β-sheet between the Cas6 thumb domain and the thumb of Cas7.1 (Fig. 1h, j). Another change lies in Cas8/5^{HB}, which disassociates from Cas6 and tilts slightly to accommodate TniQ (Fig. 1j).

Classifications of particle images also revealed a figure-eight shaped supercomplex that accounts for 5% of particles (Supplementary information, Figs. S1e and S2). The two-fold symmetry of TniQ dimer implies another binding site for an additional VcCascade. However, modeling a VcCascade to the second binding site results in severe clash between two VcCascade (Supplementary information, Fig. S7a), suggesting that one TniQ dimer is incapable of inducing VcCascade dimerization. Structural modeling revealed that the supercomplex is composed of two head-to-head arranged VcCascade in each side, and two side-by-side arranged TniQ dimers in the center (Fig. 1k). No significant conformational changes are observed in VcCascade-TniQ in the context of the supercomplex, except for a slight tilt in Cas8/5^{HB} (Supplementary information, Fig. S7b). Surprisingly, two extra globular densities attached to each of Cas8/5^{HB} are observed. We interpret the density as the Cas6-crRNA hairpin complex due to their structural resemblance (Fig. 1k, l). Structural analysis suggests that the interaction mode between Cas6-crRNA hairpin and Cas8/5^{HB} in the supercomplex is the same as in the TniQ-free VcCascade (Supplementary information, Fig. S7c, d). A dimer of VcCascade is not observed in our dataset or, to our knowledge, described in previous studies, indicating that formation of the supercomplex is TniQ dependent.

In summary, our atomic structure of VcCascade-TniQ reveals the mechanism of TniQ recruitment to VcCascade and provides a structural glimpse into the transposon-associated CRISPR-Cas system. TniQ dimer binds to the PAM distal end of VcCascade through an interaction network involving Cas6 and Cas8/5^{HB}, consistent with recent reports.[1,12] The observation of the VcCascade-TniQ supercomplex implies that high-order structure of VcCascade may play a role in DNA transposition. Further work is required to investigate how other components are recruited by VcCascade-TniQ. The mechanistic insights into this RNA-guided DNA insertion process would facilitate the manipulation of this system for genome editing and therapeutic applications in the future.

Cryo-EM maps of VcCascade-TniQ, VcCascade, and VcCascade-TniQ supercomplex were deposited in the EMDB under accession numbers EMD-20908, EMD-20909, and EMD-20910, respectively. The atomic model of VcCascade-TniQ was deposited in the RCSB PDB under the accession number 6UVN.

ACKNOWLEDGEMENTS

We thank Thomas Klose and Valorie Bowman for help with cryo-EM, Steven Wilson for computation, and Ramaswamy Subramanian for helpful discussions. This work made use of the Purdue Cryo-EM Facility and facilities in the Bindley Bioscience Center. L.C. is supported by the Department of Biological Sciences at Purdue University and a Showalter Trust Research Award.

AUTHOR CONTRIBUTIONS

L.C. supervised the study. H.Z., R.X. and Z.L. prepared samples. Z.L., H.Z. and L.C. collected and processed cryo-EM data. Z.L., H.Z. and R.X. performed biochemical analysis. All authors analyzed the data. Z.L. and L.C. prepared the manuscript with input from H.Z. and R.X.

ADDITIONAL INFORMATION

Supplementary information accompanies this paper at https://doi.org/10.1038/s41422-019-0268-y.

Competing interests: The authors declare no competing interests.

Zhuang Li[1], Heng Zhang[1], Renjian Xiao[1] and Leifu Chang[1,2]
[1]Department of Biological Sciences, Purdue University, West Lafayette, IN 47907, USA and [2]Purdue University Center for Cancer Research, West Lafayette, IN 47907, USA
These authors contributed equally: Zhuang Li, Heng Zhang
Correspondence: Leifu Chang (lchang18@purdue.edu)

REFERENCES

1. Klompe, S. E., Vo, P. L. H., Halpin-Healy, T. S. & Sternberg, S. H. *Nature* **571**, 219–225 (2019).
2. Strecker, J. et al. *Science* **365**, 48–53 (2019).
3. Peters, J. E., Makarova, K. S., Shmakov, S. & Koonin, E. V. *Proc. Natl Acad. Sci. USA* **114**, E7358–E7366 (2017).
4. Faure, G. et al. *Nat. Rev. Microbiol.* **17**, 513–525 (2019).
5. Chowdhury, S. et al. *Cell* **169**, 47–57 (2017).
6. Guo, T. W. et al. *Cell* **171**, 414–426 (2017).
7. Peng, R. et al. *Cell Res.* **27**, 853–864 (2017).
8. Rollins, M. F. et al. *Mol. Cell* **74**, 132–142 (2019).
9. Mitra, R., McKenzie, G. J., Yi, L., Lee, C. A. & Craig, N. L. *Mob. DNA* **1**, 18 (2010).
10. Laity, J. H., Lee, B. M. & Wright, P. E. *Curr. Opin. Struct. Biol.* **11**, 39–46 (2001).
11. Schumacher, M. A. et al. *Nat. Commun.* **9**, 1139 (2018).
12. Halpin-Healy, T. S., Klompe, S. E., Sternberg, S. H. & Fernández, I. S. *bioRxiv*. https://doi.org/10.1101/706143 (2019).