nature communications

Article

https://doi.org/10.1038/s41467-024-55716-7

Structural basis for RNA-guided DNA degradation by Cas5-HNH/Cascade complex

Received: 11 June 2024

Accepted: 19 December 2024

Published online: 04 February 2025

Check for updates

Yanan Liu[1,5], Lin Wang[1,5], Qian Zhang[2,5], Pengyu Fu[3,5], Lingling Zhang[2], Ying Yu[3], Heng Zhang[2,3] ✉ & Hongtao Zhu[1,4] ✉

Type I-E CRISPR (clustered regularly interspaced short palindromic repeats)–Cas (CRISPR-associated proteins) system is one of the most extensively studied RNA-guided adaptive immune systems in prokaryotes, providing defense against foreign genetic elements. Unlike the previously characterized Cas3 nuclease, which exhibits progressive DNA cleavage in the typical type I-E system, a recently identified HNH-comprising Cascade system enables precise DNA cleavage. Here, we present several near-atomic cryo-electron microscopy (cryo-EM) structures of the *Candidatus Cloacimonetes* bacterium Cas5-HNH/Cascade complex, both in its DNA-bound and unbound states. Our analysis reveals extensive interactions between the HNH domain and adjacent subunits, including Cas6 and Cas11, with mutations in these key interactions significantly impairing enzymatic activity. Upon DNA binding, the Cas5-HNH/Cascade complex adopts a more compact conformation, with subunits converging toward the center of nuclease, leading to its activation. Notably, we also find that divalent ions such as zinc, cobalt, and nickel down-regulate enzyme activity by destabilizing the Cascade complex. Together, these findings offer structural insights into the assembly and activation of the Cas5-HNH/Cascade complex.

Prokaryotes have evolved an adaptive immune system known as CRISPR (clustered regularly interspaced short palindromic repeats)–Cas (CRISPR-associated proteins) to defend against foreign genetic elements, such as invading bacteriophage and plasmids[1–3]. The CRISPR array consists of conserved repeat sequences interspersed with variable spacer sequences, which are derived from alien nucleic acids[1,4–6]. Transcripts from CRISPR loci are processed into short, mature CRISPR RNAs (crRNAs), which assemble with Cas proteins to form large ribonucleoprotein surveillance complexes[7,8]. The coevolution of bacteria and invading genetic elements has led to remarkable diversification of CRISPR-Cas systems, which are now classified into two major classes: Class 1 and Class 2[9,10]. These classes are further divided into six types (I–VI) and numerous subtypes (e.g., type I-A to I-F)[11,12]. Class 1 systems, which are the most prevalent among bacteria and archaea, featured multiple protein effectors[13], while Class 2 systems rely on a single effector protein, such as Cas9, Cas12 or Cas13, which has made them widely used in genome-editing technologies[14–17].

The canonical Cascade (CRISPR-associated complex for antiviral defense) surveillance complex in type I-E systems is composed of five Cas proteins: Cas5, Cas6, Cas7, Cas8 and Cas11—arranged in unequal stoichiometric amounts[18–22]. The type I-E immune response occurs in three stages: spacer acquisition, crRNA processing, and interference[23,24]. During spacer acquisition, the Cas1 and Cas2 proteins, which form a butterfly-shaped dimer, capture short foreign genetic elements and integrate them into the host CRISPR array as spacers, enabling bacteria to acquire new immune memories[25–27]. In the second stage, the endoribonuclease Cas6 processes pre-crRNA into mature crRNAs, each containing a 5′ handle, a spacer, and a 3′ stem-loop structure about

[1]Beijing National Laboratory for Condensed Matter Physics, Institute of Physics, Chinese Academy of Sciences, Beijing, China. [2]Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Tianjin Medical University, Tianjin, China. [3]Department of Pharmacology, School of Basic Medical Sciences, Tianjin Medical University, Tianjin, China. [4]University of Chinese Academy of Sciences, Beijing, China. [5]These authors contributed equally: Yanan Liu, Lin Wang, Qian Zhang, Pengyu Fu. ✉e-mail: zhangheng134@gmail.com; hongtao.zhu@iphy.ac.cn

61 nucleotides long[28,29]. Cas6 remains bound to the 3′ stem-loop, facilitating the recruitment of other Cascade subunits to form a complete complex[30,31]. In the interference stage, the Cascade complex recognized invasive DNA by identifying a protospacer-adjacent motif (PAM) and a sequence complementary to the crRNA[32,33]. Cas3, a helicase-nuclease protein, is then recruited to degrade the foreign DNA[34,35], though it is not a stable component of the Cascade complex[36,37].

Recently, a variant of the type I-E system lacking Cas3 was discovered in the *Candidatus Cloacimonetes* bacterium through advanced bioinformatics techniques employing a fast locality-sensitive hashing-based clustering (FLSHclust) algorithm[38]. In this variant, an HNH nuclease domain replaces the function of Cas3 and is fused to the C-terminus of Cas5[38].

In this study, we present near-atomic resolution cryo-electron microscopy (cryo-EM) structures of the Cas5-HNH/Cascade complex in both DNA-bound and unbound states. In the DNA-bound states, we obtain one structure where the target strand is inserted into the HNH domain, though the density for the non-target strand is unclear, with only 6-7 base pairs of double-stranded DNA visible near the PAM recognition site, likely due to its flexibility. Upon binding to target DNA, the Cas5-HNH/Cascade complex adopts a more compact conformation than in the unbound state. The 5′-end of the target DNA makes a pronounced U-turn, aligning parallel to the β-strands of Cas6 and intruding into the HNH domain, consistent with the structure reported by Zhang et al.[39] Interestingly, when bound to double-stranded DNA, the Cas5-HNH complex cleaves the target strand first, followed by the non-target strand—opposite to the cleavage order observed in the RuvC domain of Cas12a[40–43]. Finally, we find that the complex is sensitive to several divalent cations, including cesium, cobalt, nickel, and zinc, which primarily destabilize the Cascade complex. These findings provide insights into regulating and activating the HNH nuclease domain, which operates differently from previously characterized type I CRISPR-Cas systems.

Results

Biochemistry and structure determination of Cas5-HNH/Cascade complex

The wild-type Cas5-HNH/Cascade complexes were recombinantly expressed in *E. coli*, with a 10XHis tag on Cas8. SDS-PAGE analysis confirmed the co-purification of all required subunits, including Cas5-HNH, Cas6, Cas7, and Cas11, along with Cas8 (Fig. 1a, Supplementary Fig. 1a, b). Prior to the cryo-EM studies, digestion assays were performed to validate enzymatic activity, showing that the wild-type Cas5-HNH/Cascade complex efficiently cleaves target DNA (Supplementary Fig. 1c). Three conformations of the target-free complex were resolved: two intact complexes (apo_Conf1, apo_Conf2) and one partially assembled complex (apo_Conf3), with resolutions of 2.47 Å, 3.10 Å and 2.55 Å, respectively (Supplementary Figs. 2, 3).

Previous studies have demonstrated that the Cas5-HNH/Cascade complex can cleave both single-stranded DNA (ssDNA) and double-stranded DNA (dsDNA)[38]. To further investigate the interaction of target DNA with the complex, fluorescence-detection size-exclusion chromatography (FSEC[44]) was used. A fluorophore group, FAM, was attached to the 3′ end of the target strand of dsDNA (TD), while Cy3 was labeled at the 3′ end of the non-target strand of dsDNA (NTD). FSEC data showed peak migration for both TD and NTD, indicating that the Cas5-HNH/Cascade complex targets both ssDNA and dsDNA (Supplementary Fig. 4). To capture stable structures bound to target DNA, we incubated a catalytic mutant ribonucleoprotein (RNP) complex with target ssDNA or dsDNA, followed by cryo-EM analysis.

For the ssDNA-bound complex, two conformations were observed: an intact complex (ssDNA_Conf1) and a partially assembled complex (ssDNA_Conf2), with resolutions of 3.00 Å and 2.81 Å, respectively (Supplementary Figs. 5–8). Additionally, we collected 10,379 micrographs to resolve the structure of the complex bound to dsDNA, achieving a resolution of 2.93 Å with partial density of the NTD (dsDNA_Conf, Fig.1, Supplementary Fig. 9).

Overall architecture of the Cas5-HNH/Cascade complex

The cryo-EM structures of Cas5-HNH/Cascade complex exhibit a distinctive seahorse-shaped architecture. Take dsDNA_Conf for example, the complex measures ~130 Å in width and 178 Å in length (Fig. 1b), which is comparable to canonical Cascade complexes in type I-E CRISPR systems[18–22]. In contrast, the apo_Conf1 structure is about 185 Å in length, indicating a 5–7 Å extension compared to the target-bound states (Supplementary Fig. 2a). This suggests that the complex adopts a more compact conformation upon binding to target DNA. In the partially assembled complexes, such as apo_Conf3, key components like the HNH domain of Cas5, the N terminal domain of Cas6, Cas8, and Cas11 were unresolved. Due to the absence of Cas8, apo_Conf3 measures around 150 Å in height, making it significantly shorter than Conf1 and Conf2 (Supplementary Fig. 2a–c).

In the intact complexes, six Cas7 subunits (labeled as Cas7.1 to Cas7.6) are arranged along the crRNA to form a helical filament, which serves as the structural scaffold of Cas5-HNH/Cascade complex (Supplementary Fig. 10a). Despite their similar overall conformations, subtle differences were observed among the Cas7 subunits (Supplementary Fig. 10b). For instance, a β-hairpin in Cas7.1 (residues 185-210) and several α-helices in Cas7.6 (residues 55-137 and 149-165) were missing in the cryo-EM map (Supplementary Fig. 10b). The two ends of the Cas7 filament are capped by Cas8 and Cas6, while Cas5-HNH and Cas11 lie center of the seahorse-shaped structure (Fig. 1b–g). The 3′ crRNA handle adopts a hook-like shape, clamping onto Cas6 RNase, while the 5′ end is shielded by Cas5 (Supplementary Fig. 10c, d). In addition to interactions with Cas7, the crRNA establishes extensive contacts with Cas5, Cas6 and Cas8 via hydrogen bonds and π-π interactions (Supplementary Fig. 11). For example, W160 and Y231 of Cas5 engage in π-π interactions with −7G and −5G, respectively, while R56 and K167 form hydrogen bonds with the crRNA (Supplementary Fig. 11).

Upon binding with target DNA, the 5′-end of the target strand undergoes a pronounced U-turn, aligning parallel to the β-strands of Cas6, and intruding into the HNH domain (Fig. 1i). For the non-target strand, dsDNA-like density spanning 6–7 base pairs (5′-end of NTD) near the PAM recognition site of Cas8 was identified, though the remaining NTD density was not observed (Fig. 1c,f), likely due to flexibility.

HNH domain functions as a nuclease in Cas5-HNH Cascade

Previous studies have identified three highly conserved residues in the canonical HNH motif: two histidines and one asparagine[45,46]. In our structures, the first conserved histidine (H310) in the HNH domain of Cas5 (referred to as HNH^{Cas5}) is positioned at the end of β-strand 1 (Fig. 2a). However, unlike the classic HNH motif, where the second active site residue is typically an asparagine, this position is replaced by aspartate (D324) in HNH^{Cas5}. Sequence alignment of the type I-E Cas5-HNH family confirms that this substitution is conserved (Supplementary Fig. 12a, b), suggesting that aspartate plays a crucial role in regulating HNH^{Cas5} activity. Biochemical analysis revealed that alanine substitutions at H310 and D324 completely abolished cleavage activity (Fig. 2b). Furthermore, electrophoretic mobility shift assay (EMSA) demonstrated that the D324A and H333A mutant did not affect DNA binding (Supplementary Fig. 12c). D324N mutant also attenuated the efficiency of digestion (Supplementary Fig. 12d).

The classic Cascade complex contains two subtype-specific Cas11 subunits in its core[18–22]. However, in the Cas5-HNH/Cascade complex, only one Cas11 subunit is present (Fig. 1, Supplementary Fig. 10f). Instead, an HNH nuclease domain is fused to the C terminus of Cas5 (HNH^{Cas5}), which consists of two β-sheets and three α-helices (Fig. 2a), occupying the position typically held by the second Cas11 subunit. Since both Cas3 and HNH^{Cas5} function as nucleases, Cas3 is dispensable

https://doi.org/10.1038/s41467-024-55716-7

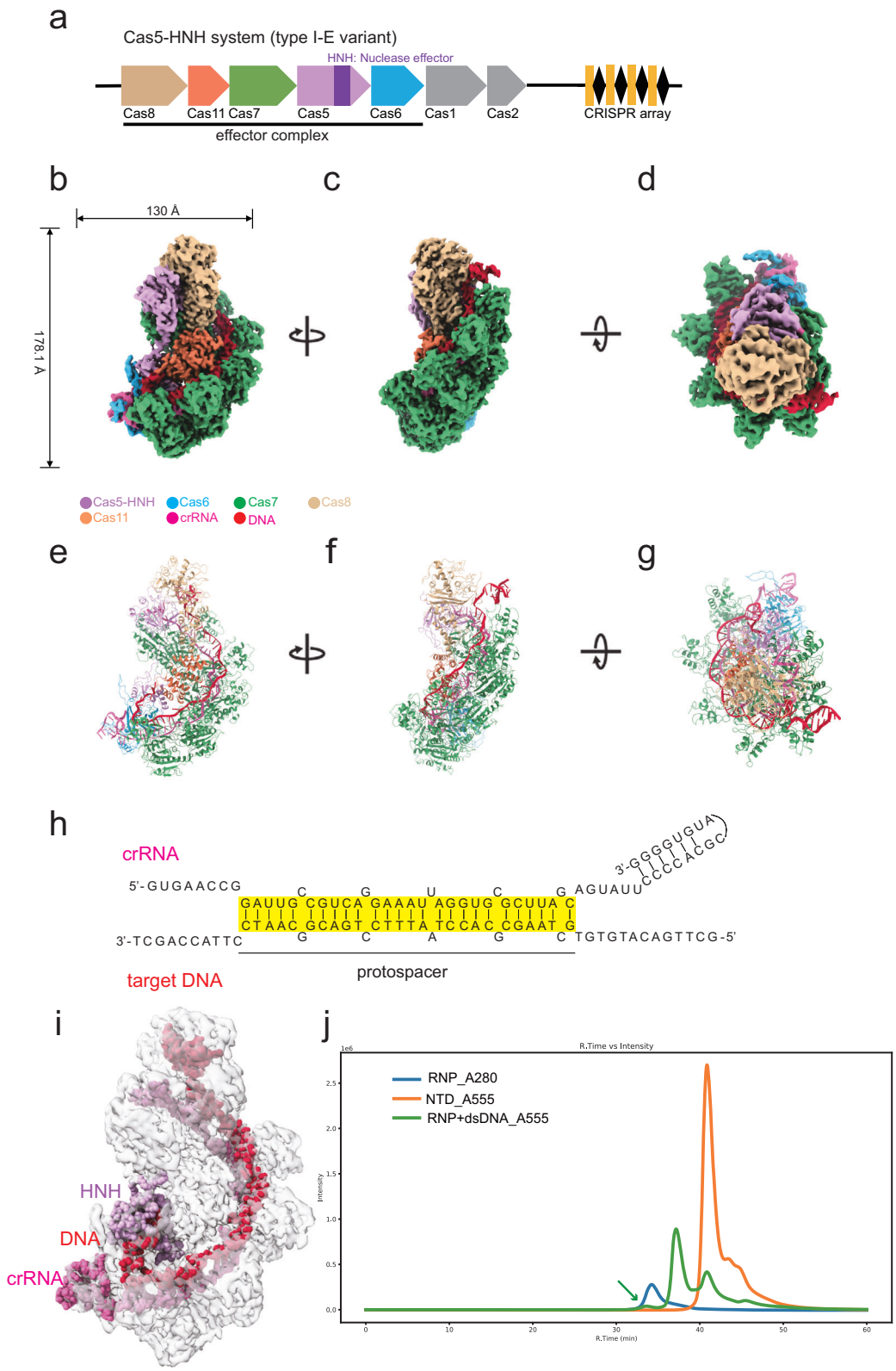

Fig. 1 | Cryo-EM structures of the Cas5-HNH/Cascade complex in Type I-E CRISPR-Cas system. a Locus diagrams of the Cas5-HNH system. The Cas genes encoding the full-length Cas5-HNH (Uniprot ID: A0A1V6F8C5), Cas6 (Uniprot ID: A0A1V6F8C4), Cas7 (Uniprot ID: A0A1V6F8B5), Cas11 (Uniprot ID: A0A1V6F8C9) and Cas8 (Uniprot ID: A0A1V6F8D1) proteins of the type I-E Cas5-HNH system are derived from *Candidatus Cloacimonetes* bacterium ADurb.Bin088. Cryo-EM map of the Cas5-HNH/Cascade complex displayed in side views (**b**, **c**) and top-down view (**d**). Six copies of Cas7 are shown in green. The main lobe of Cas5 and the HNH domain are highlighted in plum. Cas6, Cas8 and Cas11 are colored in blue, burlywood and coral, respectively. The crRNA is represented in hot pink and the dsDNA is colored in crimson. **e**–**g** Cartoon representations of the Cas5-HNH/Cascade complex structure, corresponding to views in (**b**–**d**). **h** Schematic representation of the crRNA and target DNA. **i** The crRNA (hot pink) and DNA (crimson) are shown, with the Cas5-HNH/Cascade complex displayed in surface representation. **j** FSEC trace indicating that RNP binds to non-target DNA labeled by Cy3. The arrow highlights the peak shift of DNA (Source data are provided as a Source Data file).

https://doi.org/10.1038/s41467-024-55716-7

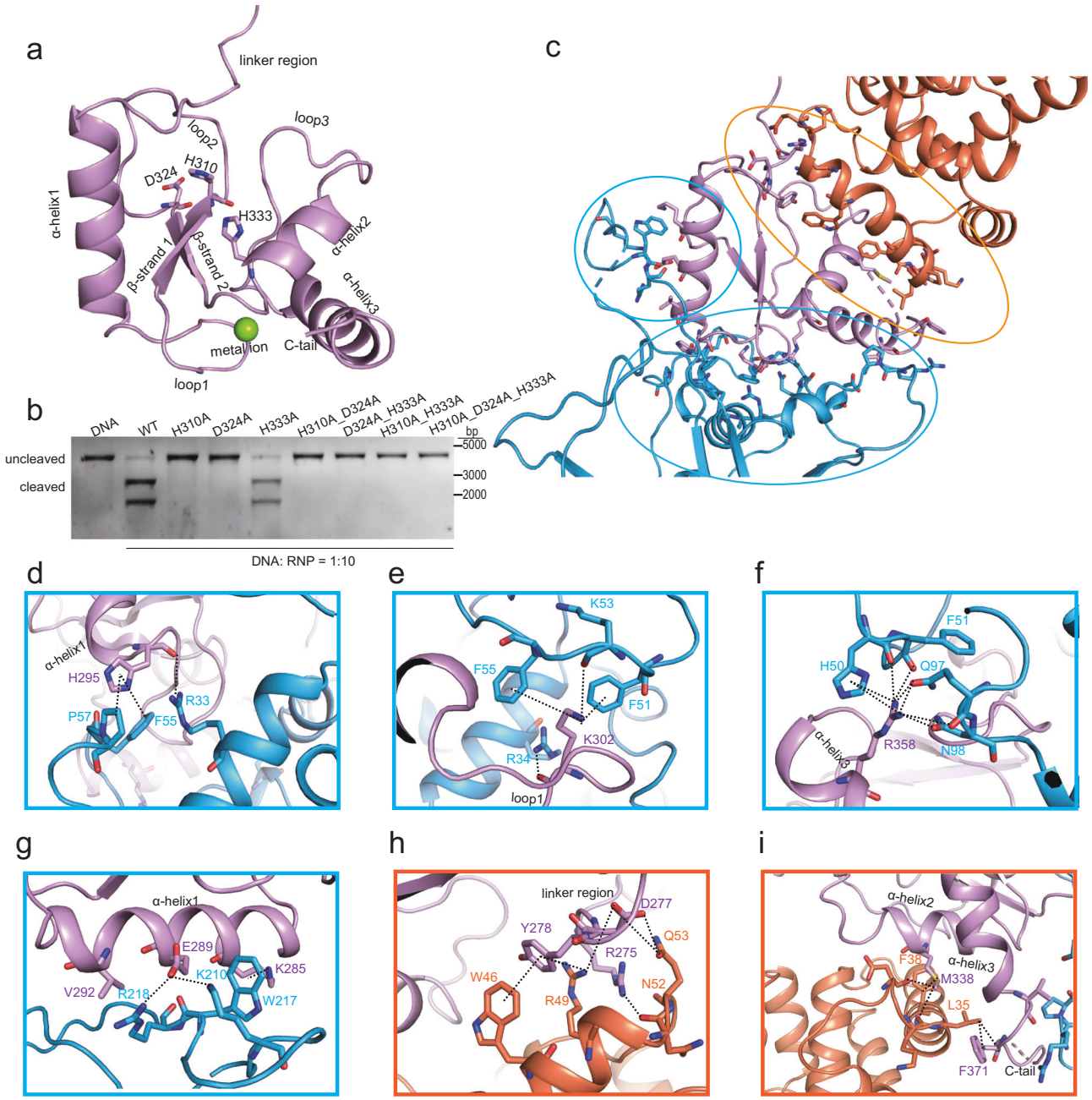

Fig. 2 | HNH domain functions as an endonuclease in type I-E variant Cas5-HNH /Cascade complex. a Structure of the HNH domain at the C-terminus of Cas5. HNH^{Cas5} consists of three α-helices, two β-strands and several loops. Three featured residues (H310, D324, H333) are depicted as sticks, color-coded by elements. The divalent metal ion is represented by a green sphere. **b** Plasmid digestion analysis results for the corresponding mutants. **c**–**i** Interactions between the HNH domain and Cas6/Cas11, highlighted by dashed lines. The residues of Cas5, Cas6, and Cas11 are shown in plum, blue and coral, respectively.

during the interference stage of bacterial immune response due to the presence of the HNH domain in the type I-E variant system.

Structural comparisons using the DALI server revealed that the Cas5-HNH domain shares strong structural similarity with the HNH endonuclease from the deep-sea thermophilic bacteriophage GVE2 (PDB: 5h0m)[47], *Pseudomonas syringae* Cap5 (PDB: 8FMF)[48], and the type IV Coupling Complex (T4CC) from *Legionella pneumophila* (PDB: 6SZ9)[49], with z-scores of 8.6, 6.4, and 6.1, respectively. The HNH domain of type II nuclease Cas9 (PDB: 6JDQ)[50] also exhibits structural similarity to Cas5-HNH, though with a lower z-score of 2.7 (Supplementary Fig. 13a, b). Despite this structural resemblance, these proteins display low sequence similarity (Supplementary Fig. 13c).

Interactions involving the HNH domain are critical for regulating the function of the complex. Investigation of the interfaces between HNH^{Cas5} and Cas6/Cas11 revealed extensive interactions (Fig. 2c). Hydrogen bonds were observed involving H295, K302, and R358 in HNH^{Cas5}, R33, K53, R34, Q97, F51 and N98 in Cas6 (Fig. 2d–f). Additionally, cation-π interactions were identified at the HNH-Cas6 interface, involving $K302^{Cas5}$-$F51^{Cas6}$, $R358^{Cas5}$-$H50^{Cas6}$ and $K285^{Cas5}$-$W217^{Cas6}$, (Fig. 2e–g). Salt bridges between $E289^{Cas5}$ and $R218^{Cas6}$/$K210^{Cas6}$ further stabilize this interaction (Fig. 2g). To confirm the functional significance of these interactions, we designed several mutants and assessed their enzymatic efficiency. As expected, some of mutations affecting the HNH-Cas6 interface resulted in a significant reduction in

https://doi.org/10.1038/s41467-024-55716-7

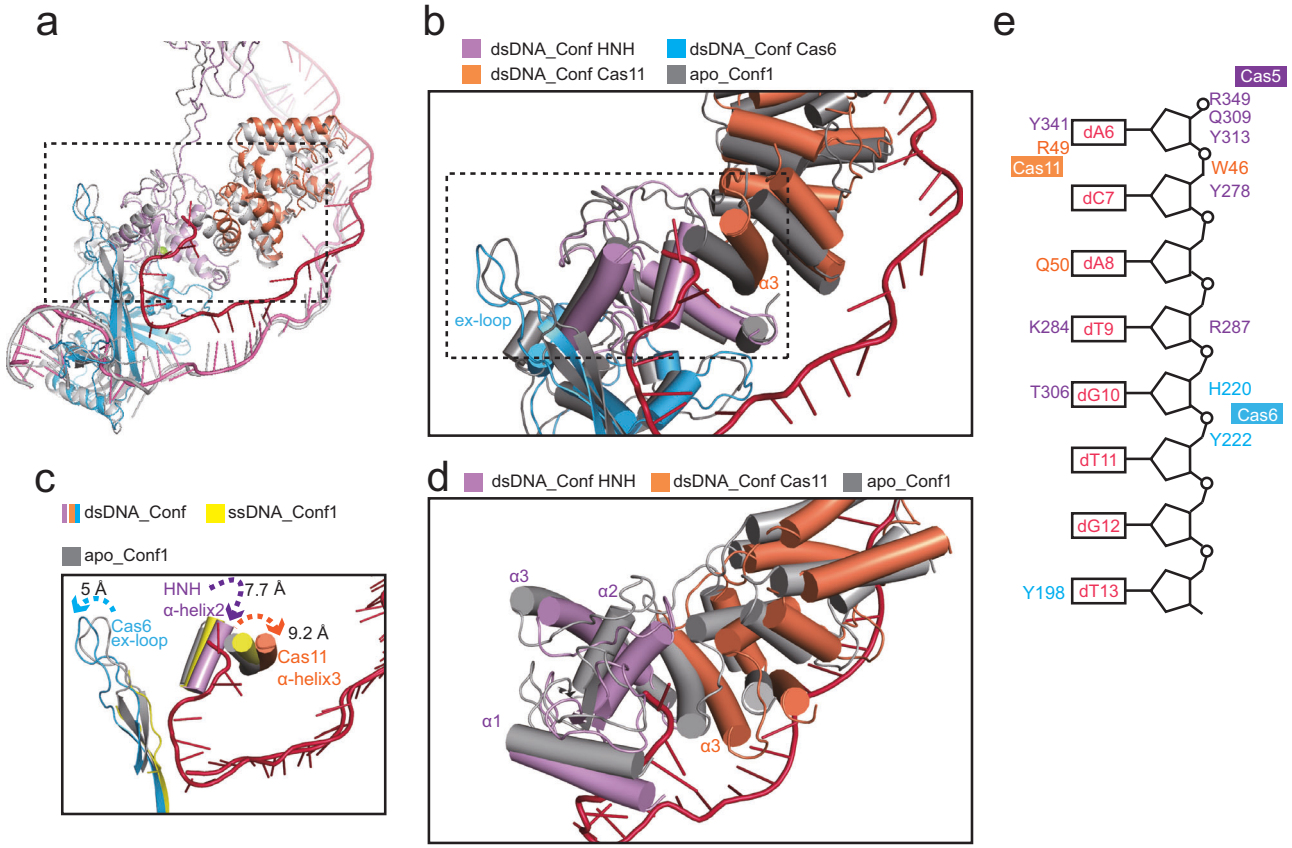

Fig. 3 | Dynamics of Cas5-HNH /Cascade complex upon binding target DNA. a–d Movement of the HNH domain, Cas11, and Cas6. Cas6 and Cas11 shift away from the target DNA, while the HNH domain moves closer to the DNA. **e** Residues of the complex involved in interactions with the 5′-end of DNA.

DNA cleavage efficiency (Supplementary Fig. 14a–c). Notably, the E289A (HNH) mutation disrupted target DNA binding (Supplementary Fig. 14e). Furthermore, combined mutations at M338, V369 and F371, along with neighboring residues in Cas6 (D16, R17, P18), impaired the function of the complex (Supplementary Fig. 14b). The HNH domain is connected to the N-terminal domain of Cas5 via a long linker (Supplementary Fig. 10d), prompting us to investigate whether residues in this linker regulate nuclease activity. A combined mutation of R275A, D277A, and Y278A in the linker region resulted in a slight reduction in nuclease activity (Supplementary Fig. 14d). Interestingly, these residues are associated with interaction with Cas11 (Fig. 2h), and mutating the interacting partners (W46, R49, and Q53) of Cas11 completely abolished the enzyme's activity (Supplementary Fig. 14d). M338 and F371, which interact with Cas11 through hydrophobic interactions, also modestly influence nuclease activity (Fig. 2i, Supplementary Fig. 14b). Taken together, our structural and biochemical data demonstrated that interactions between Cas6/Cas11 and HNH^{Cas5} are essential for the proper function of Cas5-HNH/Cascade complex.

To elucidate the cleavage mechanism of Cas5-HNH, we compared the dsDNA-bound state with the apo state. Significant movements were observed among Cas5, Cas6, and Cas11. Upon DNA binding, Cas6 and Cas11 shift away from the 5′ end, opening the active channel for cleavage, while the HNH domain moves toward the substrate (Fig. 3a–d). The shift of the extension loop (residues 199-225, termed as ex-loop) between the two β-strands of Cas6 measures 5 Å, while the movements of the HNH domain and Cas11 range from 8 to 10 Å (Fig. 3c). Notably, α-helix 3 (α3) of Cas11 undergoes a twisting conformational change upon DNA binding, with the C-terminus of Cas11-α3 moving ~9 Å away from the DNA while the N-terminus remains nearly in its original position (Fig. 3d). Additionally, several residues that interact with the target DNA, such as Y278 (Cas5) and W46 (Cas11), are also involved in the interaction surface between subunits (Figs. 2h and 3e).

As anticipated, mutations in residues associated with the 5′-end of DNA significantly reduced nuclease activity, with specific mutations at Y278, T306, and Q309 in the HNH domain nearly abolishing the enzyme's function (Supplementary Fig. 14i). Notably, most of these mutations did not affect DNA-binding efficiency; some mutants, such as K284A and Q309A, even exhibited increased DNA affinity (Supplementary Fig. 14j).

Given the unique dsDNA-cleavage mechanism of Cas5-HNH, which differs from type II CRISPR-Cas systems that require a RuvC domain to cleave the non-target strand[38,40–42], we conducted cleavage order experiments. The results indicated that NTD cleavage occurs following TD digestion (Supplementary Fig. 15a, b). The non-target strand was cleaved less efficiently than the target strand (Supplementary Fig. 15c). Additionally, we evaluated the mismatch tolerance of the Cas5-HNH/Cascade complex using various mismatched DNA substrates. Mismatches at PAM-distal positions (29-32) did not affect Cas5-HNH nuclease activity, whereas mismatches at PAM-proximal positions (17-19) significantly impaired its tolerance (Supplementary Fig. 16).

Comparison of the structures reveals distinct conformational differences

A comparison between the target-bound complex and the unbound state reveals significant conformational differences (Fig. 4). Notably, dsDNA_Conf, ssDNA_Conf1 and apo_Conf2 exhibit more compact conformations compared to apo_Conf1 (Fig. 4a–c). The global RMSD values between apo_Conf1 and ssDNA_Conf1, as well as dsDNA_Conf are 2.27 Å and 2.81 Å, respectively.

https://doi.org/10.1038/s41467-024-55716-7

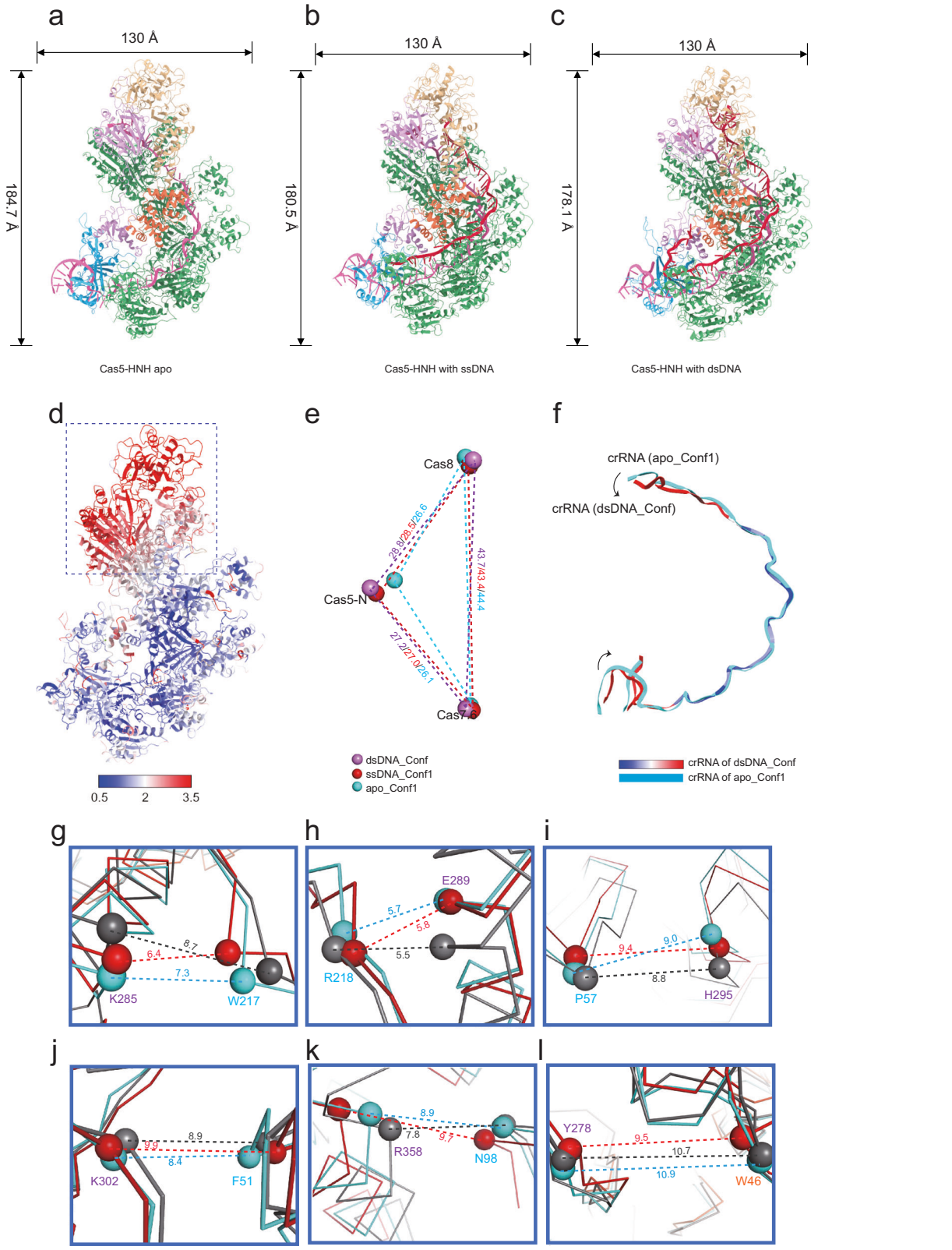

Fig. 4 | Comparison of the structures unbound or bound with DNA reveals distinct conformational differences. Cartoon representations of the Cas5-HNH/Cascade complex in its apo_Conf1 (**a**), ssDNA_Conf1 (**b**), and dsDNA_Conf (**c**) states. Heights and widths are denoted in Å. **d** RMSD distribution calculated between ssDNA_Conf1 and apo_Conf1. The most significant differences occur in the Cas5, Cas8 and Cas7.6 subunits, highlighted by the dashed rectangle. **e** Schematic illustrating the neighboring distances (Å) between the center of mass (COM) of Cas8, Cas7.6 and N-terminal domain (1-270) of Cas5. **f** Superposition of the HNH domain from dsDNA_Conf (color-coded by RMSD values) and apo_Conf1 (cyan). The RMSD color matched that in (**d**). **g**–**l** Distances between the Cα atoms of representative amino acids in HNH^{Cas5}. The atoms from ssDNA_Conf1, apo_Conf1, and apo_Conf2 are colored by red, cyan, and gray, respectively.

https://doi.org/10.1038/s41467-024-55716-7

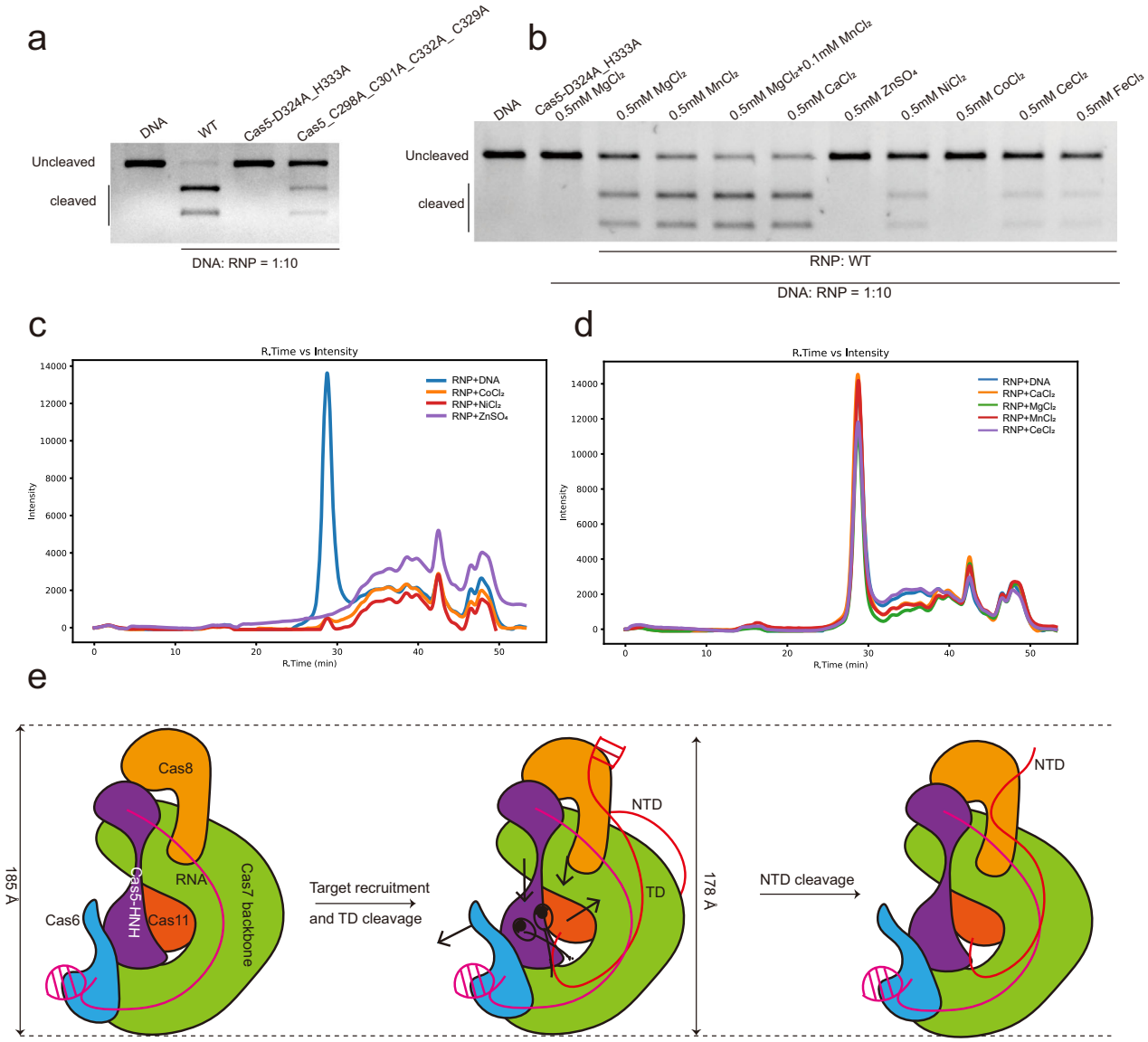

Fig. 5 | Divalent ions play important roles in regulating the Cas5-HNH/Cascade complex and proposed model. a Biochemistry analysis shows that mutating these four cysteines abolished both DNA binding and cleavage activity. **b** Enzyme activity assays in the presence of various metal ions. **c** FSEC traces of the Cas5-HNH/Cascade complex with Co^{2+}, Ni^{2+} and Zn^{2+}. **d** FSEC traces for the complex in the presence of 0.5 mM Mg^{2+}, Mn^{2+} and Ca^{2+}. The wavelength of the FSEC detector is set to 280 nm. **e** A schematic model illustrates the dynamic changes occurring in the complex following target recruitment.

To investigate the conformational changes induced by TD binding, we aligned the crRNA in apo_Conf1 and ssDNA_Conf1, tracking the shifts in the center of mass (COM) of the associated subunits (Supplementary Fig. 17a, b). Structural overlays between apo_Conf1 and ssDNA_Conf1 highlight that the primary differences occur in the Cas5, Cas8 and Cas7.6 subunits (Fig. 4d, e). Significant conformational variations are also observed in the crRNA across these structures. An overlay of crRNA from apo_Conf1 and dsDNA_Conf reveals a clear shift at both ends of the crRNA, with a calculated RMSD of 4.12 Å (Fig. 4f).

Previous studies have demonstrated that activation of the CRISPR Cascade complex involves subunit movement towards the nuclease center[20,22,35]. Consistently, upon ssDNA binding, the N-termini of Cas5, Cas7.6, and Cas8 moved towards the HNH^{Cas5}. The calculated RMSD between ssDNA_Conf1 and apo_Conf1 for these three subunits is ~3.50 Å.

We further measured the distances between the Cα atoms of key amino acids on HNH^{Cas5} and their interaction partners (Fig. 4g–l). In ssDNA_Conf1, some distances between key amino acids are reduced compared to those in apo_Conf1 (Fig. 4g, l), while others show increased distances (Fig. 4h–k).

Divalent ions in regulating the Cas5-HNH/Cascade complex

Sequence alignment analysis reveals that the HNH^{Cas5} is conserved across several species (Supplementary Fig. 12). Four conserved cysteines, C298, C301, C329 and C332, form a zinc-finger like structure in HNH^{Cas5} (Supplementary Fig. 12e), as previous reported[51,52]. Mutations of these cysteines abolished both the binding to and digestion of the target DNA (Fig. 5a).

Previous research has established that the HNH nuclease contains one or two metal-binding sites[51,53], typically for zinc and magnesium[51]. We tested various metal ions, including magnesium, manganese, calcium, cesium, cobalt, nickel and zinc, to determine their effects on the enzyme efficiency of the Cas5-HNH/Cascade complex. The findings demonstrated that, in the presence of 0.5 mM magnesium, manganese and calcium, the enzymatic efficiency of the Cas5-HNH/Cascade complex was comparable across these ions (Fig. 5b). However, adding

cesium, cobalt, nickel and zinc significantly reduced DNA digestion efficiency. Especially for zinc and cobalt, the activity of the enzyme was nearly abolished (Fig. 5b).

Interestingly, in FSEC[39] experiments conducted in the presence of these metal ions, we observed that the addition of cobalt, nickel and zinc led to the disassociation of the complex (Fig. 5c), which may explain the observed reduction in enzymatic efficiency. Although cesium also decreased enzyme efficiency, it did not induce dissociation of the complex (Fig. 5d).

Discussion

In this study, we elucidated near-atomic structures of the recently discovered Cas5-HNH Cascade complex. Our structures show that the HNH^{Cas5}, which cleaves substrate DNA, replaced the Cas3 nuclease in the classic type I-E CRISPR system. We revealed that the HNH^{Cas5} forms extensive interactions with adjacent subunits, including Cas6 and Cas11, and that specific amino acids at the HNH^{Cas5} interface can influence enzyme activity. Structural analysis indicates that upon binding to target DNA, the Cas5-HNH/Cascade complex adopts a more compact conformation, with subunits Cas5, Cas7.6, and Cas8 moving toward the nuclease center to activate the enzyme. Of note, the 5′-end of the target DNA makes a pronounced U-turn, aligning parallel to the β-strands of Cas6 and intruding into the HNH domain, consistent with the structure reported by Zhang et al.[39] Mutants associated with substrate binding in HNH domain significantly impact cleavage activity. Cas6 and Cas11 move away from the substrate, while HNH domain shift toward DNA, suggesting its central role for nuclease. The target-free structure (apo_Conf1) aligns perfectly with structure 8YB6 (Supplementary Fig. 18a, b). In contrast, when compared to 8YHA, the Cas11 of our dsDNA-bound structure exhibits a noticeable displacement away from the nuclease center (Supplementary Fig. 18c, d). We also uncovered that the complex is vulnerable to divalent ions, including cobalt, nickel and zinc, which can cause disassociation and consequently reduce enzymatic efficiency. Upon binding to dsDNA, the target strand of dsDNA (TD) is cleaved first, followed by NTD. In summary, we present cryo-EM structures and biochemical analyses of the substrate-bound Cas5-HNH/Cascade complex, shedding light on the dymanic change around the nuclease center (Fig. 5e). This study paves the way for further molecular and structural investigations into the intriguing functions of this variant type I-E CRISPR-Cas system.

Methods

Protein expression and purification

The Cas genes encoding full-length Cas5-HNH (Uniprot ID: A0A1V6F8C5), Cas6 (Uniprot ID: A0A1V6F8C4), Cas7 (Uniprot ID: A0A1V6F8B5), Cas11 (Uniprot ID: A0A1V6F8C9) and Cas8 (Uniprot ID: A0A1V6F8D1) proteins of the type I-E Cas5-HNH system were sourced from *Candidatus Cloacimonetes* bacterium ADurb.Bin088. All genes were codon optimized and synthesized by GENEWIZ. The Cas5-HNH and N-terminal His_{10}-tagged Cas8 were fused into a pCDFDuet vector, while the genes for Cas7 and Cas6 were cloned into the pRSFDuet-1 vector. Additionally, a gene encoding a 61-nt crRNA was inserted into a pACYCT2 vector. All plasmids were co-transformed into Escherichia coli BL21 (DE3) cells, and protein expression was induced with 0.2 mM isopropyl-β-D-thiogalactoside when the optical density at 600 nm (OD600 nm) reached 0.6. The cells were incubated for 12–16 h at 20 °C. The harvested cells were resuspended in the binding buffer (25 mM Tris-HCl pH 7.5, 500 mM NaCl, 3 mM β-mercaptoethanol, 5 mM imidazole, and 1% glycerol). Following sonication and centrifugation, the lysate was loaded onto Ni-NTA resin. The resin was washed three times with the binding buffer and eluted with the same buffer supplemented with 300 mM imidazole. The eluate was then applied to a HiTrap Heparin HP column (Cytiva), and the peak fractions containing the target protein complex were concentrated and further purified using a Superdex 6 column (Cytiva) equilibrated with running buffer (25 mM Tris-HCl pH 7.5, 200 mM NaCl, and 2 mM DTT). The peak fractions were collected and analyzed by SDS-PAGE.

The RNP was incubated with target DNA labeled with fluorophore groups and subsequently loaded onto a fluorescence size exclusion chromatography (FSEC) system[44] with a Superose 6 column to assess binding efficiency. The Cy5- and Cy3-labeled DNA substrates were synthesized by Tsingke Biotech.

Cryo-EM sample preparation and data collection

To prepare the cryo-EM sample, the protein was concentrated to 1 mg/mL. The Vitrobot Mark IV (FEI) was then used to freeze the cryo-EM grids. A 3.2 μL sample was applied to the plasma cleaned GIG R1.2/1.3 holey gold grids, using a blot force of 2 for 2 s in 100% humidity. The grids were subsequently plunged into liquid ethane cooled by liquid nitrogen. Cryo-EM data were collected using a 300 kV Titan Krios equipped with a Falcon4 direct electron detector (Gatan). Additional collection parameters are provided in Supplementary Table 1.

Protein stability determination

Prior to loading the sample onto the FSEC[44] system, 3 μg of protein was incubated on ice for 30 min with 0.5 mM solutions of various metal ions, including $MgCl_2$, $MnCl_2$, $CaCl_2$, $ZnSO_4$, $CoCl_2$, $CeCl_2$, and $NiCl_2$. The samples were then loaded to the FSEC system equipped with a Superose 6 column, and the absorbance was measured at 280 nm.

Cryo-EM data analysis

The cryo-EM data analysis was conducted using cryoSPARC[54], unless otherwise noted. For the dataset of the Cas5-HNH/Cascade complex unbound to DNA, a total of 7382 movies were recorded. Beam-induced motion was corrected using MotionCor2[55]. Subsequently, defocus values were estimated through CTF patch estimation[56]. Protein particles were picked using Blob picker with the minimum and maximum diameters set to 100 Å and 220 Å, respectively. Following rounds of 2D classification, 1336 k good particles were selected from a total of 1989 k raw particles. To obtain the initial model, six classes were defined during the ab-initio reconstruction, and models exhibiting strong protein features were selected for a further heterogeneous refinement. Notably, the raw particles generated by Blob picker were utilized in this round of heterogeneous refinement. Four classes were chosen for an additional round of heterogeneous refinement, with class1 and class4 being selected for subsequent analysis. For class1, one round of heterogeneous refinement was performed to eliminate junk particles. To mitigate the effects of preferred orientation, a particle rebalance was performed with a rebalance factor of 1.0, followed by one round of non-uniform refinement. The remaining 20,320 particles were then converted to Relion[57] format using Pyem[58], and spIsonet[59] was employed to perform misalignment and anisotropy correction (apo_Conf1). For class4, a round of heterogeneous refinement was conducted, and the class with the highest resolution was selected for a final run of non-uniform refinement (apo_Conf2). The resolution for apo_Conf1 and apo_Conf2, as estimated by cryoSPARC[54], were 2.47 Å and 3.10 Å, respectively. DeepEMhancer[60] was used to sharpen both maps. For the dataset of the Cas5-HNH/Cascade complex bound to target ssDNA and dsDNA, a total of 4040 and 10,379 micrographs were collected respectively. The data processing was similar to that of the apo-complex, except that apo_Conf1 was used as a template for particle picking. Additional data are provided in Table S1.

Model building

Homology models for Cas5-HNH, Cas6, Cas7, Cas8 and Cas11 were generated by Alphafold[61]. USCF-Chimera[62] was utilized to fit the predicted models into the corresponding density map, and all models were subsequently combined into a single model. The chain IDs were re-assigned in Coot[63]. Manually adjustments were made in Coot based on high-quality features. Several residues, specifically residues 1-80 in

chain C (Cas8), residues 54-138, 148-166 and 314-322 in Chain G (Cas7) and residues 190-208 and 256-263 in chain K (Cas7), were removed due to obscured densities. The RNA structure was manually built in Coot[63], and real space refinement in Phenix[64] was performed to refine the model against the cryo-EM map.

We observed significant conformational differences between apo_Conf1 and apo_Conf2, making it impractical to directly fit the entire apo_Conf1 model into the apo_Conf2 cryo-EM map. To address this, the chains from apo_Conf1 model were saved individually in Chimera[62]. These saved chains were then reloaded into Chimera[62] and fitting was performed chain by chain into the apo_Conf2 cryo-EM map. The chains were combined in Chimera[62] before proceeding to Coot[63] for further manual adjustments. Residues with poor densities, including residues 1-80 in chain C (Cas8), residues 54-138 and 195-204 in Chain G (Cas7), residues 91-97 in chain F (Cas7), residues 190-208 and 256-233 in chain K (Cas7) and residues 197-205 in chain J (Cas7), were removed due to the poor densities. Five rounds of real space refinement in Phenix[64] were performed to further refine the model. Data collection and model refinement statistics are summarized in Table S1. A similar procedure was followed to build the model ssDNA_Conf1 and ssDNA_Conf2, where the apo_Conf1 structure was fitted to the ssDNA bound maps in Chimera, followed by manual refinement in Coot[63]. Phenix[64] was then employed to refine the model against the corresponding maps. The models of complex binding to DNA were built in similar way.

In vitro DNA cleavage assay

The target DNA sequence was inserted into a pET-derived vector, and FspI (NEB) was used to linearize the resulting plasmids for the target DNA cleavage assay. The Cas5-HNH/Cascade complex and target DNA substrate were pre-diluted in a cleavage buffer containing 50 mM Tris-HCl, pH 7.5, 100 mM NaCl, 5 mM $MgCl_2$, and 1 mM DTT. To assess the cleavage activity of the Cas5-HNH system, the target DNA was incubated with either the wide-type or mutant complex at molar ratios ranging from 1:10 to 1:50. The reaction was maintained at 37 °C for 1 h. Following incubation, 40 mM EDTA and 1 mg/ml protease K were added to terminate the reaction. Cleavage products were detected using 0.5% TBE agarose gel stained with StarStain Red Nucleic Acid Dye (GenStar).

Electrophoretic mobility shift assay (EMSA)

The binding of target DNA to the Cas5-HNH/Cascade complex was assessed using DNA EMSA. The TD and 5′-Cy3 labeled NTD were annealed by heating to 95 °C for 3 min, followed by gradual cooling to 25 °C. The purified Cas5-HNH/Cascade complex was incubated with the annealed dsDNA at 37 °C for 30 min in a buffer containing 50 mM Tris-HCl (pH 7.5), 10 mM NaCl, and 1 mM DTT, with increasing concentrations of wide-type or mutant Cas5-HNH complex. Samples were then loaded onto a 5% native PAGE gel and run for 30 min at 120 V and 4 °C. The reaction products were visualized using a Tanon 5200 imaging system.

Statistics and reproducibility

In vitro, DNA cleavage assay and EMSA were performed at least three times independently.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The coordinates and cryo-EM maps have been deposited in the Electron Microscopy Data Bank under accession codes EMD-61880 (dsDNA_Conf), EMD-60328 (ssDNA_Conf1), EMD-60330 (ssDNA_Conf2), EMD-60233 (apo_Conf1), EMD-60235 (apo_Conf2), and EMD-60297 (apo_Conf3). The coordinates also have been deposited in the Protein Data Bank under accession codes 9JXS (dsDNA_Conf), 8ZP7 (ssDNA_Conf1), 8ZP9 (ssDNA_Conf2), 8ZLU (apo_Conf1), 8ZM3 (apo_Conf2) and 8ZOL (apo_Conf3), respectively. Source data are provided with this paper.

References

1. Barrangou, R. et al. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712 (2007).
2. van der Oost, J., Westra, E. R., Jackson, R. N. & Wiedenheft, B. Unravelling the structural and mechanistic basis of CRISPR-Cas systems. *Nat. Rev. Microbiol.* **12**, 479–492 (2014).
3. Wiedenheft, B., Sternberg, S. H. & Doudna, J. A. RNA-guided genetic silencing systems in bacteria and archaea. *Nature* **482**, 331–338 (2012).
4. Pourcel, C., Salvignol, G. & Vergnaud, G. CRISPR elements in Yersinia pestis acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151**, 653–663 (2005).
5. Garneau, J. E. et al. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71 (2010).
6. Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151**, 2551–2561 (2005).
7. Brouns, S. J. et al. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964 (2008).
8. Xue, C. & Sashital, D. G. Mechanisms of type I-E and I-F CRISPR-Cas systems in Enterobacteriaceae. *EcoSal Plus* **8**, https://doi.org/10.1128/ecosalplus.ESP-0008-2018 (2019).
9. Makarova, K. S. et al. Evolutionary classification of CRISPR-Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* **18**, 67–83 (2020).
10. Murugan, K., Babu, K., Sundaresan, R., Rajan, R. & Sashital, D. G. The revolution continues: newly discovered systems expand the CRISPR-Cas toolkit. *Mol. cell* **68**, 15–25 (2017).
11. Koonin, E. V., Makarova, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.* **37**, 67–78 (2017).
12. Makarova, K. S. et al. An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* **13**, 722–736 (2015).
13. Makarova, K. S., Zhang, F. & Koonin, E. V. SnapShot: class 1 CRISPR-Cas systems. *Cell* **168**, 946–946 e941 (2017).
14. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. USA* **109**, E2579–E2586 (2012).
15. Abudayyeh, O. O. et al. RNA targeting with CRISPR-Cas13. *Nature* **550**, 280–284 (2017).
16. Jiang, F. et al. Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science* **351**, 867–871 (2016).
17. Wang, S. W. et al. Current applications and future perspective of CRISPR/Cas9 gene editing in cancer. *Mol. Cancer* **21**, 57 (2022).
18. Zhao, H. et al. Crystal structure of the RNA-guided immune surveillance Cascade complex in Escherichia coli. *Nature* **515**, 147–150 (2014).
19. Hale, C. R. et al. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139**, 945–956 (2009).
20. Hayes, R. P. et al. Structural basis for promiscuous PAM recognition in type I-E Cascade from E. coli. *Nature* **530**, 499–503 (2016).
21. Mulepati, S., Heroux, A. & Bailey, S. Structural biology. Crystal structure of a CRISPR RNA-guided surveillance complex bound to a ssDNA target. *Science* **345**, 1479–1484 (2014).
22. van Erp, P. B. et al. Mechanism of CRISPR-RNA guided recognition of DNA targets in Escherichia coli. *Nucleic Acids Res.* **43**, 8381–8391 (2015).

23. Liu, T. Y. & Doudna, J. A. Chemistry of Class 1 CRISPR-Cas effectors: binding, editing, and regulation. *J. Biol. Chem.* **295**, 14473–14487 (2020).
24. Nunez, J. K., Lee, A. S., Engelman, A. & Doudna, J. A. Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature* **519**, 193–198 (2015).
25. Nunez, J. K. et al. Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat. Struct. Mol. Biol.* **21**, 528–534 (2014).
26. Rollins, M. F. et al. Cas1 and the Csy complex are opposing regulators of Cas2/3 nuclease activity. *Proc. Natl. Acad. Sci. USA* **114**, E5113–E5121 (2017).
27. Yoganand, K. N., Sivathanu, R., Nimkar, S. & Anand, B. Asymmetric positioning of Cas1-2 complex and Integration Host Factor induced DNA bending guide the unidirectional homing of protospacer in CRISPR-Cas type I-E system. *Nucleic acids Res.* **45**, 367–381 (2017).
28. Gesner, E. M., Schellenberg, M. J., Garside, E. L., George, M. M. & Macmillan, A. M. Recognition and maturation of effector RNAs in a CRISPR interference pathway. *Nat. Struct. Mol. Biol.* **18**, 688–692 (2011).
29. Jore, M. M. et al. Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat. Struct. Mol. Biol.* **18**, 529–536 (2011).
30. Luo, M. L. et al. The CRISPR RNA-guided surveillance complex in Escherichia coli accommodates extended RNA spacers. *Nucleic Acids Res.* **44**, 7385–7394 (2016).
31. Sokolowski, R. D., Graham, S. & White, M. F. Cas6 specificity and CRISPR RNA loading in a complex CRISPR-Cas system. *Nucleic Acids Res.* **42**, 6532–6541 (2014).
32. Mojica, F. J. M., Diez-Villasenor, C., Garcia-Martinez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155**, 733–740 (2009).
33. Reeks, J., Naismith, J. H. & White, M. F. CRISPR interference: a structural perspective. *Biochem. J.* **453**, 155–166 (2013).
34. Yoshimi, K. et al. Dynamic mechanisms of CRISPR interference by Escherichia coli CRISPR-Cas3. *Nat. Commun.* **13**, 4917 (2022).
35. Xiao, Y., Luo, M., Dolan, A. E., Liao, M. & Ke, A. Structure basis for RNA-guided DNA degradation by Cascade and Cas3. *Science* **361**, https://doi.org/10.1126/science.aat0839 (2018).
36. Bondy-Denomy, J. & Davidson, A. R. To acquire or resist: the complex biological effects of CRISPR-Cas systems. *Trends Microbiol.* **22**, 218–225 (2014).
37. Csorgo, B. et al. A compact Cascade-Cas3 system for targeted genome engineering. *Nat. Methods* **17**, 1183–1190 (2020).
38. Altae-Tran, H. et al. Uncovering the functional diversity of rare CRISPR-Cas systems with deep terascale clustering. *Science* **382**, eadi1910 (2023).
39. Zhang, C. et al. Mechanisms for HNH-mediated target DNA cleavage in type I CRISPR-Cas systems. *Mol. Cell* **84**, 3141–3153 e3145 (2024).
40. Jeon, Y. et al. Direct observation of DNA target searching and cleavage by CRISPR-Cas12a. *Nat. Commun.* **9**, 2777 (2018).
41. Strohkendl, I., Saifuddin, F. A., Rybarski, J. R., Finkelstein, I. J. & Russell, R. Kinetic basis for DNA target specificity of CRISPR-Cas12a. *Mol. Cell* **71**, 816–824 e813 (2018).
42. Stella, S. et al. Conformational activation promotes CRISPR-Cas12a catalysis and resetting of the endonuclease activity. *Cell* **175**, 1856–1871 e1821 (2018).
43. Swarts, D. C. & Jinek, M. Cas9 versus Cas12a/Cpf1: Structure-function comparisons and implications for genome editing. *Wiley Interdiscip. Rev. RNA* **9**, e1481 (2018).
44. Hattori, M., Hibbs, R. E. & Gouaux, E. A fluorescence-detection size-exclusion chromatography-based thermostability assay for membrane protein precrystallization screening. *Structure* **20**, 1293–1299 (2012).
45. Jinek, M. et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821 (2012).
46. Xu, S. Y. & Gupta, Y. K. Natural zinc ribbon HNH endonucleases and engineered zinc finger nicking endonuclease. *Nucleic Acids Res.* **41**, 378–390 (2013).
47. Zhang, L. et al. Structural and functional characterization of deep-sea thermophilic bacteriophage GVE2 HNH endonuclease. *Sci. Rep.* **7**, 42542 (2017).
48. Rechkoblit, O. et al. Activation of CBASS Cap5 endonuclease immune effector by cyclic nucleotides. *Nat. Struct. Mol. Biol.* **31**, 767–776 (2024).
49. Meir, A. et al. Mechanism of effector capture and delivery by the type IV secretion system from Legionella pneumophila. *Nat. Commun.* **11**, 2864 (2020).
50. Sun, W. et al. Structures of Neisseria meningitidis Cas9 Complexes in Catalytically Poised and Anti-CRISPR-Inhibited States. *Mol. Cell* **76**, 938–952 e935 (2019).
51. Bujnicki, J. M., Radlinska, M. & Rychlewski, L. Atomic model of the 5-methylcytosine-specific restriction enzyme McrA reveals an atypical zinc finger and structural similarity to betabetaalphaMe endonucleases. *Mol. Microbiol.* **37**, 1280–1281 (2000).
52. Mehta, P., Katta, K. & Krishnaswamy, S. HNH family subclassification leads to identification of commonality in the His-Me endonuclease superfamily. *Protein Sci.* **13**, 295–300 (2004).
53. Mate, M. J. & Kleanthous, C. Structure-based analysis of the metal-dependent mechanism of H-N-H endonucleases. *J. Biol. Chem.* **279**, 34763–34769 (2004).
54. Punjani, A., Rubinstein, J. L., Fleet, D. J. & Brubaker, M. A. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
55. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
56. Zhang, K. Gctf: Real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
57. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
58. Schorb, M., Haberbosch, I., Hagen, W. J. H., Schwab, Y. & Mastronarde, D. N. Software tools for automated transmission electron microscopy. *Nat. methods* **16**, 471–477 (2019).
59. Liu, Y. T., Hu, J. & Zhou, Z. H. Resolving the preferred orientation problem in CryoEM reconstruction with self-supervised deep learning. *Microsc. Microanal.* **29**, 1918–1919 (2023).
60. Sanchez-Garcia, R. et al. DeepEMhancer: a deep learning solution for cryo-EM volume post-processing. *Commun. Biol.* **4**, 874 (2021).
61. Hryc, C. F. & Baker, M. L. AlphaFold2 and CryoEM: Revisiting CryoEM modeling in near-atomic resolution density maps. *iScience* **25**, 104496 (2022).
62. Pettersen, E. F. et al. UCSF Chimera-a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
63. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. Sect. D. Biol. Crystallogr.* **60**, 2126–2132 (2004).
64. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. Sect. D. Struct. Biol.* **74**, 531–544 (2018).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (32471015), Chinese Academy of Sciences (E2VK311RA1) to H. Zhu, the Natural Science Foundation of Tianjin Municipal Science and Technology Commission (23JCZDJC00410), and National Natural Science Foundation of China (32322040) to H. Zhang. We thank Z. Q. Guo and H.

S. Li from Shuimu Biosciences for the cryo-EM data collection. We thank X. Li and Q. He for assistance with experiments.

Author contributions

Conceptualization: H.Z. and H. Zhu; Protein purification and Biochemical assays: Q.Z., P.Y.F., L.L.Z.; EM data collection and processing: Y.L., L.W., and H. Zhu; Structure modeling: Y.L., L.W., and H. Zhu; Supervision: Y. Yu, H.Z., and H. Zhu; Manuscript writing: Y.L., L.W., H.Z., and H. Zhu.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-024-55716-7.

Correspondence and requests for materials should be addressed to Heng Zhang or Hongtao Zhu.

Peer review information *Nature Communications* thanks David Taylor and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.