



Mechanisms for target recognition and cleavage by the Cas12i RNA-guided endonuclease

Heng Zhang ^{1,3}✉, Zhuang Li ^{1,3}, Renjian Xiao ^{1,3} and Leifu Chang ^{1,2}✉

Cas12i is a recently identified type V CRISPR-Cas endonuclease that predominantly cleaves the non-target strand of a double-stranded DNA substrate. This nicking activity of Cas12i could potentially be used for genome editing with high specificity. To elucidate its mechanisms for target recognition and cleavage, we determined cryo-EM structures of Cas12i in multiple functional states. Cas12i pre-orders a seven-nucleotide seed sequence of the crRNA for target recognition and undergoes a two-step activation through crRNA-DNA hybridization. Formation of 14 base pairs activates the nickase activity, and 28-bp hybridization promotes cleavage of the target strand. The atomic structures and mechanistic insights gained should facilitate the manipulation of Cas12i for genome editing applications.

C RISPR-Cas systems provide adaptive immunity in bacteria and archaea to resist invasion of mobile genetic elements (MGEs) via RNA-guided cleavage of target nucleic acids^{1–6}. CRISPR-Cas acquired immunity occurs through three major steps: incorporation of short fragments of invasive nucleic acids into the CRISPR array as new spacers by the Cas1–Cas2 integrase complex (step 1, acquisition or adaptation); transcription and maturation of crRNAs (CRISPR RNAs), which assemble with Cas proteins into a ribonucleoprotein (RNP) complex (step 2, expression step); recognition of the invasive nucleic acids and cleavage by the RNP complex (step 3, interference step)^{7,8}. Dynamics between the host and MGEs have resulted in diverse CRISPR-Cas systems, which are grouped into class 1 systems (types I, III and IV) and class 2 systems (types II, V and VI)^{9–12}. Class 2 systems utilize a single effector endonuclease and have been harnessed for genome editing and targeting tools^{13–16}, with Cas9 from the type II systems being the most widely used^{13,14}. In addition to Cas9, three type V CRISPR-Cas systems have been successfully harnessed for genome editing in mammalian cells, including Cas12a (Cpf1)¹⁷, Cas12b (C2c1)^{18,19} and Cas12e (CasX)²⁰. Both type II and type V effectors adopt a similar bilobed architecture comprising a recognition lobe (REC) and a nuclease lobe (NUC). Cas9 endonucleases employ both the RuvC and HNH domains for DNA cleavage; however, Cas12 endonucleases use a single RuvC domain for DNA cleavage.

Recently, a number of new type V systems have been identified, revealing functional diversity in the type V systems²¹ and providing new opportunities to improve and expand CRISPR-based applications. Among the newly identified type V endonucleases, Cas12i is a single RNA-guided nickase that predominantly cleaves the non-target strand (NTS) of target DNA with a 5'-TTN-3' protospacer adjacent motif (PAM). The nicking activity of Cas12i can be potentially used for genome editing with high specificity²¹. Like Cas12a, Cas12i can process its own pre-crRNA into mature crRNA, providing a promising alternate platform for multiplexed genome editing²¹. Intriguingly, the protein sequences of Cas12i effectors are evolutionarily diverged from other sub-types in the type V systems²¹. The absence of an acquisition module further suggests the specialized features of the CRISPR-Cas12i system²¹. Together, sequence and functional differences between Cas12i and other type

V endonucleases indicate that Cas12i may have different molecular mechanisms in target recognition and cleavage.

Structure of Cas12i-crRNA-dsDNA

To reveal the mechanism of target recognition by Cas12i, we assembled a Cas12i-crRNA-DNA ternary complex by incubating a catalytically inactive Cas12i (E894A) with a 60-nucleotide (nt) precursor crRNA (pre-crRNA), a 36-nt target DNA strand and a 10-nt non-target DNA strand containing a 5'-TTN-3' PAM sequence, and then determined the structure of the ternary complex to an average resolution of 2.9 Å using cryo-EM (Extended Data Figs. 1 and 2 and Table 1). The pre-crRNA was processed into a mature crRNA by cleavage of a segment at the 5' end by the RNase activity of Cas12i (Extended Data Fig. 1b), consistent with a previous study²¹.

The overall structure of Cas12i displays a bilobed architecture encompassing a REC composed of the REC1, REC2 and PAM-interacting (PI) domains and a NUC composed of the RuvC, Wedge (WED) and Nuc domains (Fig. 1a–g). The crRNA contains a 28-nt spacer-derived guide segment (positions 1–28) at the 3' end and a 23-nt repeat-derived sequence (positions –1 to –23) at the 5' end, with the repeat-derived sequence forming a stem-loop structure (Fig. 1h,i). The stem-loop is sandwiched by the REC2 and WED domains, with an extended U-rich sequence preceding the stem-loop held by the WED domain (Fig. 1c,e). The spacer-derived guide segment of the crRNA and the protospacer-derived sequence of the target DNA strand form a 28-bp heteroduplex, which runs through a central channel surrounded by the WED, REC1, REC2 and RuvC domains (Fig. 1d–g). The 3' end of the target strand (TS) base-pairs with the NTS, forming the PAM duplex inserted into a groove created by the PI, WED and REC1 domains (Fig. 1d).

Despite low sequence identity (5.5%) and similarity (14.2%) (Supplementary Note), a DALI search²² revealed Cas12b to be the closest structural homolog to Cas12i. Substantial differences in their overall architectures are observed (Extended Data Fig. 3a–f). First, an additional helix (α 1) is present in the WED domain of Cas12i to accommodate the extended U-rich motif of crRNA (Fig. 1e and Extended Data Fig. 3a,e). Second, the tracrRNA stem-loop region of Cas12b is replaced by two additional helices (α 3 and α 4) in Cas12i, which are involved in stabilization of the stem-loop of

¹Department of Biological Sciences, Purdue University, West Lafayette, IN, USA. ²Purdue University Center for Cancer Research, Purdue University, West Lafayette, IN, USA. ³These authors contributed equally: Heng Zhang, Zhuang Li, Renjian Xiao. ✉e-mail: zhangheng134@gmail.com; lchang18@purdue.edu

Table 1 | Statistics of cryo-EM data and structure refinement

	Cas12i ^{E894A} -crRNA-dsDNA (EMD-21541, PDB 6W5C)	Cas12i-crRNA (EMD-21551, PDB 6W62)	I1 (EMD-21552, PDB 6W64)
Data collection and processing			
Magnification	81,000	81,000	81,000
Voltage (kV)	300	300	300
Electron exposure (e ⁻ /Å ²)	54	36	36
Defocus range (μm)	1.5–2.5	1.5–2.5	1.5–2.5
Pixel size (Å)	1.05	1.05	1.05
Symmetry imposed	C1	C1	C1
Initial particle images (no.)	4,944,615	1,507,592	1,507,592
Final particle images (no.)	731,231	122,435	75,406
Map resolution (Å)	2.9	3.9	3.9
FSC threshold	0.143	0.143	0.143
Map resolution range (Å)	2.7–20	3.6–20	3.6–20
Refinement			
Initial model used	PDB 5U30	PDB 6W5C	PDB 6W5C
Model resolution (Å)	3.1	4.2	4.1
FSC threshold	0.5	0.5	0.5
Model resolution range (Å)	2.7–50	3.8–50	3.8–50
Map sharpening <i>B</i> factor (Å ²)	−91	−161	−136
Model composition			
Nonhydrogen atoms	10,186	7,390	8,506
Protein residues	1,023	910	928
Nucleotides	105	24	56
B factors (Å²)			
Protein	37.44	30.00	25.49
Nucleic acids	47.28	39.80	28.77
R.m.s. deviations			
Bond lengths (Å)	0.004	0.003	0.003
Bond angles (°)	0.625	0.617	0.568
Validation			
MolProbity score	1.90	2.26	2.03
Clashscore	8.04	16.14	9.32
Poor rotamers (%)	0.85	0.00	1.29
Ramachandran plot			
Favored (%)	92.57	90.00	92.90
Allowed (%)	7.43	9.89	7.10
Disallowed (%)	0.00	0.11	0.00

crRNA (Fig. 1e and Extended Data Fig. 3a,e). Moreover, the REC domain of Cas12i is longer than that of Cas12b by ~20 Å, featured by a long anti-parallel helix pair, α 16 and α 17, that hold the crRNA-target DNA heteroduplex (Fig. 1e and Extended Data Fig. 3b). The PI domain of Cas12i contains three additional helices (α 8, α 9 and α 12) relative to Cas12b, while the RuvC and Nuc domains of Cas12i are more compact than those of Cas12b (Extended Data Fig. 3b,d,f).

These structural differences between Cas12i and Cas12b account for their distinct functions (to be discussed in the following).

Pre-crRNA processing

Similar to Cas12a, Cas12i is capable of processing its pre-crRNA²¹. To understand the mechanism of pre-crRNA processing by Cas12i, we analyzed the structure around the 5' end of the crRNA. The first observed nucleotide at the 5' end of the crRNA in the EM map is A(−23), surrounded by H527, H528 and R22 (Fig. 2a,b). H528 is positioned close to the 5' oxygen of A(−23), whereas H527 projects away from A(−23) and possibly lies in close proximity to A(−24). To test the role of these residues, we introduced alanine substitutions at each position. Pre-crRNA cleavage assays revealed that the single mutation of H527 or H528 almost completely abolishes, while mutation of R22 reduces, pre-crRNA cleavage (Fig. 2c). W511 forms pi-stacking interactions with U(−22) and A(−23), positioning the pre-crRNA for cleavage (Fig. 2b). Mutation of W511 also reduced pre-crRNA cleavage (Fig. 2c). Thus, H527, H528 and R22 of the WED domain likely constitute the RNase catalytic site for pre-crRNA maturation, with a cleavage site possibly located between A(−23) and A(−24)²¹ (Fig. 2a and Extended Data Fig. 1b). The structural and biochemical analysis suggest that Cas12i processes pre-crRNA by an acid-base catalytic mechanism similar to those of Cas12a²³ and Cas13a²⁴, with H527 and H528 being the acid-base catalyst pair and R22 being the stabilizer. The three residues and W511 are not conserved in Cas12b (Supplementary Note), suggesting that Cas12b uses a distinct mechanism for pre-crRNA processing, such as via RNase III²⁵. Interestingly, Cas12i cleaves pre-crRNA in an extended loop preceding the stem-loop structure, unlike Cas12a, which cleaves pre-crRNA adjacent to the 5' pseudoknot²³, and Cas6, which cleaves crRNA immediately downstream of the stem-loop structure²⁶, but it is similar to Cas13a, which also cleaves a loop preceding the repeat-derived stem-loop²⁷.

PAM recognition

The PAM duplex is recognized by the PI, WED and REC1 domains (Fig. 2d). A loop from the REC1 domain and a loop from the WED domain insert into the major groove of the PAM duplex and recognize the PAM sequence via hydrogen bonds (Fig. 2d). T168 interacts with dG(−1) while H170 and S482 interact with dA(−2) of the TS. S167 forms a hydrogen bond with dC(−1*), and Y171 forms stacking interactions with dC(−1*) and dT(−2*) of the NTS. Alanine substitution of any of these residues resulted in reduced substrate cleavage activity (Fig. 2e). A loop in the PI domain is inserted into the minor groove of the PAM duplex and positioned closely to the thymine bases of dT(−2*) and dT(−3*) in the NTS (Fig. 2d); however, specific interactions were not identified due to the poor density in the region. In addition to the sequence-specific recognition, a number of positively charged residues form polar interactions with the phosphate groups of the PAM duplex (Extended Data Fig. 4).

crRNA-DNA heteroduplex recognition

The crRNA-DNA heteroduplex is clamped in the positively charged central channel of Cas12i. Both strands of 1–14 bp form extensive contacts with the WED, REC1, REC2 and RuvC domains, mainly through charged interactions between basic residues and the phosphate groups of the duplex (Extended Data Fig. 4). However, only a few residues of Cas12i contact the 15–26 bp of the heteroduplex (Extended Data Fig. 4). Single alanine substitutions for the residues (such as K483, R535 and R769) involved in recognition of the crRNA-DNA heteroduplex resulted in only modest reductions in the cleavage activity of Cas12i (Extended Data Fig. 5a–c).

Substrate cleavage

Density consistent with nucleic acid (named ‘substrate DNA’) around the RuvC active site is observed (Fig. 2f and Extended

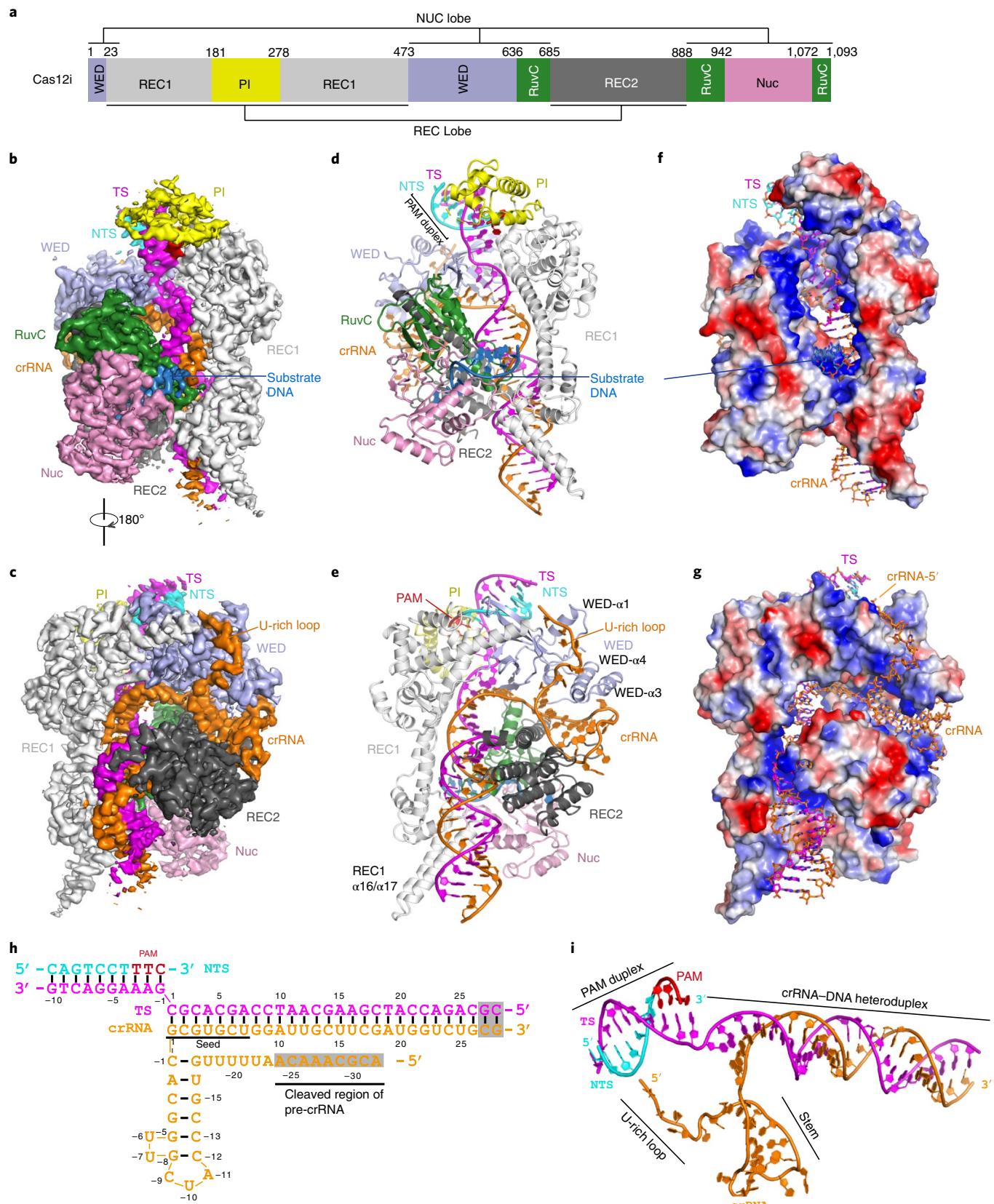


Fig. 1 | Structure of the Cas12i-crRNA-dsDNA ternary complex. **a**, Domain organization of Cas12i. **b,c**, Cryo-EM map of the Cas12i-crRNA-dsDNA ternary complex in one view (**b**) and rotated through 180° (**c**), with each domain of Cas12i color-coded as in **a**. The crRNA, TS, NTS and substrate DNA are colored in orange, magenta, cyan and sky blue, respectively. The PAM sequence is colored in red. **d,e**, Atomic models of the Cas12i-crRNA-dsDNA ternary complex in two views (as in **b** and **c**) shown in cartoon. **f,g**, Surface potentials (corresponding to the maps in **b** and **c**) of Cas12i with crRNA and DNA in stick presentation. **h**, Schematic of the crRNA and target DNA. **i**, Structure of the crRNA and target DNA in the Cas12i-crRNA-dsDNA ternary complex.

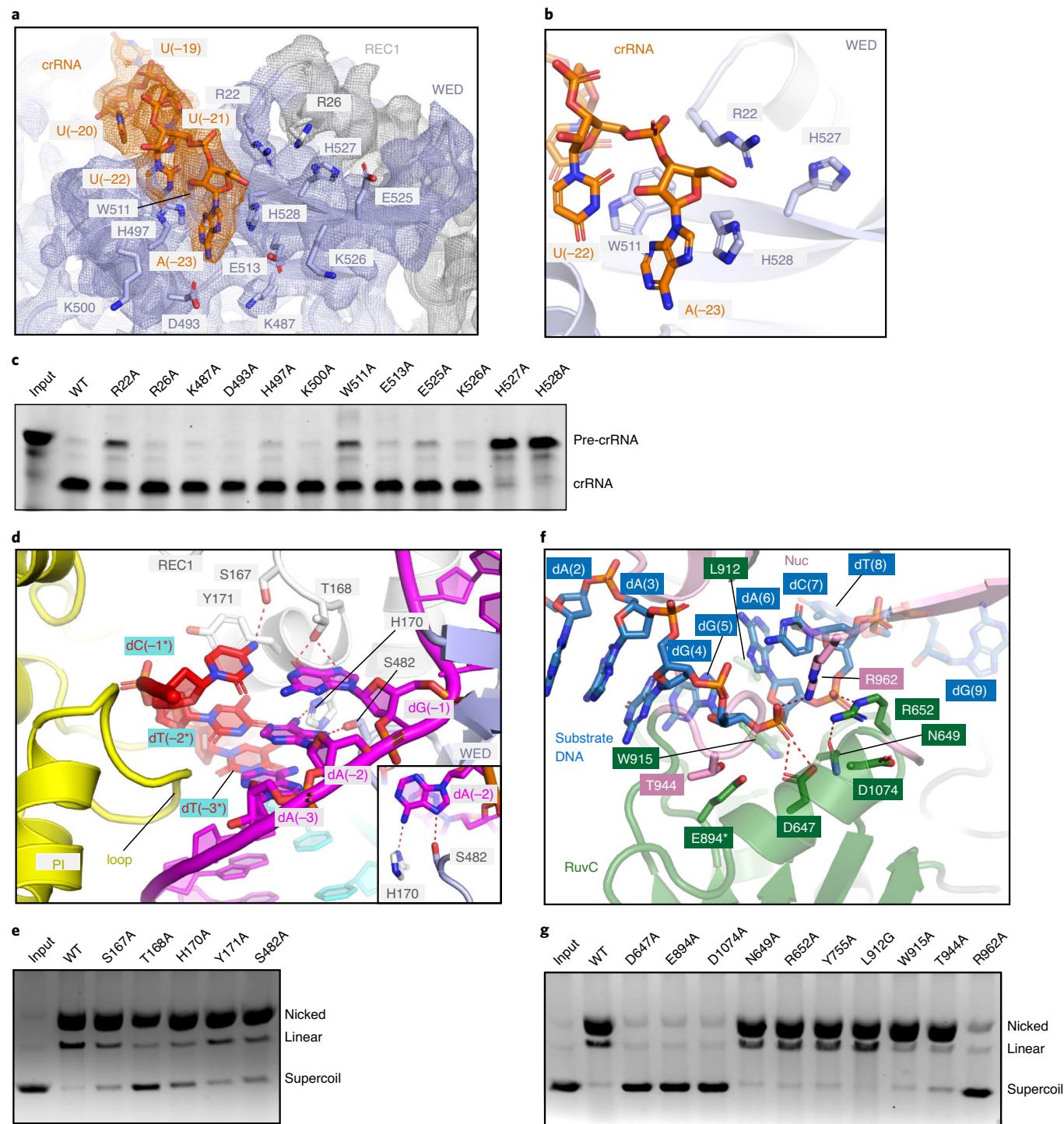


Fig. 2 | Pre-crRNA processing, substrate recognition and cleavage by Cas12i. **a**, Structure of the crRNA 5' end and surrounding Cas12i protein in the Cas12i-crRNA-dsDNA ternary complex. Polar residues potentially involved in pre-crRNA processing (tested by mutagenesis) are shown in sticks with the cryo-EM map shown in mesh. **b**, Close-up view of the 5' end of the crRNA with four key residues for pre-crRNA cleavage shown in sticks. **c**, Pre-crRNA cleavage assay using wild-type Cas12i and alanine substitutions for potential residues involved in pre-crRNA cleavage. The results shown are representative of three experiments. **d**, Detailed interactions between the PAM duplex and the REC1, WED and PI domains. Inset: the interactions between A(-2) of TS and Cas12i. **e**, Cleavage assay using a supercoil plasmid with a 5'-TTN-3' PAM sequence. The activities of wild-type Cas12i and alanine substitutions for residues involved in PAM recognition are tested. Nicking and double-strand breaks in the substrate are detected by agarose gel. The results shown are representative of three experiments. **f**, Detailed interactions between the substrate DNA and the RuvC and Nuc domains. The asterisk indicates the E894 modeled in the structure that is substituted by an alanine in the construct used for cryo-EM analysis. **g**, Substrate cleavage assay as shown in **e** using wild-type Cas12i and alanine substitutions for residues involved in substrate DNA recognition. The results shown are representative of three experiments. Uncropped images for **c**, **e** and **g** are available as source data online.

Data Fig. 2*i*), which is probably from excess DNA oligos used for Cas12i-crRNA-DNA complex assembly. Indeed, the sequence of the TS can be modeled into the density; nevertheless, unambiguous assignment was not feasible at the present resolution (Extended Data Fig. 2*i*). The presence of excess DNA in the catalytic pocket was also previously observed in Cas12b²⁸. Cas12i binds to substrate DNA mainly through residues surrounding the catalytic site in a sequence-independent manner (Extended Data Fig. 4). Consistently, Cas12i shows similar nickase activity for poly-A, -T, -C or -G substrates (Extended Data Fig. 5*d*). However, further studies will be required to investigate the substrate preference of Cas12i.

Mutagenesis and substrate cleavage assays showed that three acidic residues (D647, E894 and D1074) in the RuvC domain constitute the catalytic center (Fig. 2*g*). A basic residue (R962) from the Nuc domain points towards the catalytic center and is also critical for substrate cleavage (Fig. 2*f,g*). The structure suggests a cleavage site between dG(5) and dA(6), as the scissile phosphate between the two nucleic acid is well positioned towards the catalytic center (Fig. 2*f*). Interestingly, when W915 from the RuvC domain or T944 from the Nuc domain are substituted by alanine, cleavage of the NTS is sustained but cleavage of the TS is impaired, indicating that they play a role in TS cleavage (Fig. 2*g*).

Cas12i activation

To understand the mechanism of activation, we analyzed the structure of wild-type Cas12i in complex with crRNA and target DNA (Extended Data Fig. 6). By three-dimensional (3D) classification, we determined a 3.9-Å resolution cryo-EM structure of a Cas12i-crRNA binary complex (Extended Data Fig. 6c-f). A 7-nt guide segment (G1-U7) immediately after the crRNA hairpin is held by the WED and REC1 domains and ordered in a helical conformation (Fig. 3a and Extended Data Fig. 7a), reminiscent of the seed region observed in Cas12a and Cas12b^{23,28,29}. Furthermore, the 7-nt guide segment located in the central positively charged channel is exposed to solvent (Fig. 3a and Extended Data Fig. 7b), potentially facilitating hybridization with the TS of the substrate. Single-nucleotide mismatch in this segment resulted in significant reduction of substrate cleavage activity (Fig. 3b), indicating that substrate recognition is probably initiated from this segment. Together, the 7-nt segment is a plausible seed candidate for Cas12i.

By 3D classification, we also observed an intermediate state (I1 state) with a crRNA-DNA heteroduplex of 14–15 bp, determined to 3.9-Å resolution (Extended Data Fig. 6). A structural comparison of the binary, I1 state and Cas12i^{E894A}-crRNA-dsDNA ternary complexes shows that formation of the 14–15-bp crRNA-target DNA heteroduplex results in the outward movement of REC1 and REC2 (Fig. 3c); however, propagation of the heteroduplex to 28 bp induces an inward movement of REC1, establishing interactions between the 15–28 bp and the REC1 domain (Fig. 3d and Extended Data Fig. 4). Progression of the crRNA-DNA heteroduplex displaces a loop in the REC2 domain (aa 726–737) (Extended Data Fig. 8a-c). Deletion of this loop resulted in reduced cleavage activity (Extended Data Fig. 8d,e), possibly because the coupled conformational changes required for activation are interrupted.

Notably, a short helix (α 1) and its preceding loop in the RuvC domain (aa 897–915, the lid motif) undergo a 90° rotation upon crRNA-target DNA heteroduplex formation, and both are involved in recognition of the minor groove of the heteroduplex at positions 9–11 bp (Fig. 3e-g and Extended Data Fig. 4). Specifically, a number of polar residues contact the phosphate groups in the heteroduplex, assisting the stabilization of the heteroduplex (Extended Data Fig. 4). Rotation of the lid motif opens the substrate binding groove and thus enables substrate binding (Fig. 3e-g). Furthermore, the lid motif also stabilizes the substrate for cleavage, such as W915 described above (Fig. 2g). Collectively, the lid motif may act as a key regulator for DNA cleavage by Cas12i.

To understand the biological significance of the different lengths of crRNA-target DNA heteroduplex observed in our structures, we designed crRNAs with different lengths and measured the cleavage activity. The results suggested that a 14-nt spacer-derived sequence is sufficient to activate the nickase activity of Cas12i (Fig. 3h). However, a 28-nt spacer-derived sequence is required for cleavage of the TS, because deletion or mismatch of positions 15–28 showed no cleavage of the TS (Fig. 3b,h). Taken together, Cas12i undergoes a two-step activation, wherein formation of the 14-bp crRNA-DNA heteroduplex activates the nickase activity for cleavage of the NTS, and further base pairing to 28 bp is required for cleavage of the TS.

Discussion

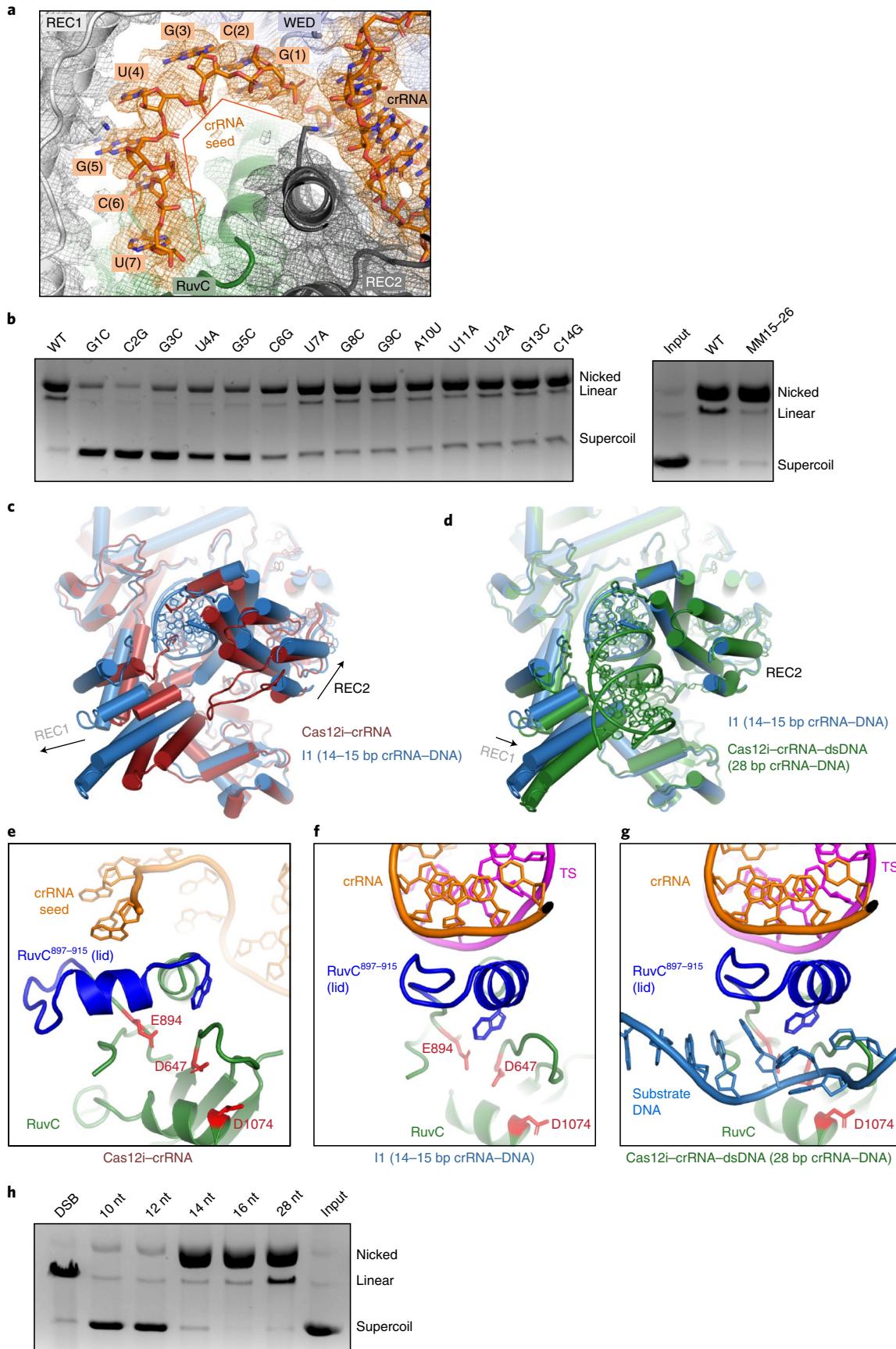
Type V CRISPR-Cas systems are featured by a single Cas12 effector protein with a RuvC endonuclease domain. So far, 11 sub-types have been reported (Cas12a-k)^{30–32}, and structures for three sub-types have been determined, including Cas12a^{23,33–41}, Cas12b^{28,29,42} and Cas12e²⁰. Most Cas12 proteins target dsDNA to create double-strand breaks. However, Cas12i predominantly cleaves the NTS, creating a nick on target dsDNA²¹.

Structure and function of Cas12i. Cas12i (as well as Cas12a) has several functional features distinct from Cas12b and Cas12e, including utilization of a single guide RNA and autonomous pre-crRNA processing. Cas12i also displays unique features in type V systems. First, Cas12i pre-orders a 7-nt seed sequence, in comparison to a 5-nt seed sequence in Cas12a^{23,35} and Cas12b^{28,29,42}. Second, Cas12i is activated by a 14–15-bp crRNA-target DNA heteroduplex, but at least 16–17 bp are required for activation of Cas12a³⁸. Another significant difference between Cas12i and other Cas12 effectors is the length of the crRNA-target DNA heteroduplex, wherein Cas12i accommodates 28 bp, compared to 20 bp for Cas12a³⁹, Cas12b²⁸ and Cas12e²⁰, as well as for Cas9^{43,44}. The presence of more nucleotides for base pairing indicates that Cas12i may impose stricter selection for a target sequence so as to create a double-strand break⁴⁵. Interestingly, formation of a 28-bp heteroduplex is critical for cleavage of the TS. The requirement of longer sequence complementation for double-strand cleavage implies that Cas12i may exhibit a higher fidelity for target selection, therefore reducing off-targeting in genome editing applications. It has been shown that rationally engineered Cas9 with a longer spacer would increase the fidelity⁴⁶, and the intrinsic preference for longer spacer base-pairing length in Cas12i should motivate the exploration of more

Fig. 3 | Mechanism of Cas12i activation. **a**, The 7-nt seed sequence in the Cas12i-crRNA binary complex, with the cryo-EM density shown in mesh. **b**, Substrate cleavage assay using wild-type crRNA and crRNAs with mismatches. Right: substrate cleavage assay using WT crRNA (WT15–26: AAGCTTACCAAGAC) and crRNA with mismatches (MMs) in positions 15–26 (MM15–26: UUCGAUGGUCUG). The results shown are representative of three experiments. **c**, Superimposition of structures for the Cas12i-crRNA binary complex and the I1 state. **d**, Superimposition of structures for the I1 state and the Cas12i-crRNA-DNA ternary complex. **e**, Enlarged view of the lid in the RuvC domain (amino acids (aa) 897–915) in the Cas12i-crRNA binary complex. Catalytic residues are colored in red. **f**, Enlarged view of the lid in the RuvC domain in the I1 state. **g**, Enlarged view of the lid in the RuvC domain in the Cas12i^{E894A}-crRNA-DNA ternary complex. **h**, Substrate cleavage assay using crRNAs with increasing lengths (10–28 nt) of spacer-derived sequence. The first lane shows the supercoil plasmid substrate treated by a restriction enzyme to create double-strand breaks (DSBs), resulting in a completely linear product. The results shown are representative of three experiments. Uncropped images for **b** and **h** are available as source data online.

Cas12i orthologs with high efficiency and accuracy. It would therefore be interesting to investigate the genome editing activity of Cas12i in eukaryotic cells.

Conserved mechanisms for Cas12 endonucleases. Similar structural patterns (stem-loop or pseudoknot) are observed in the crRNA repeat region adjacent to the spacer-derived sequence in



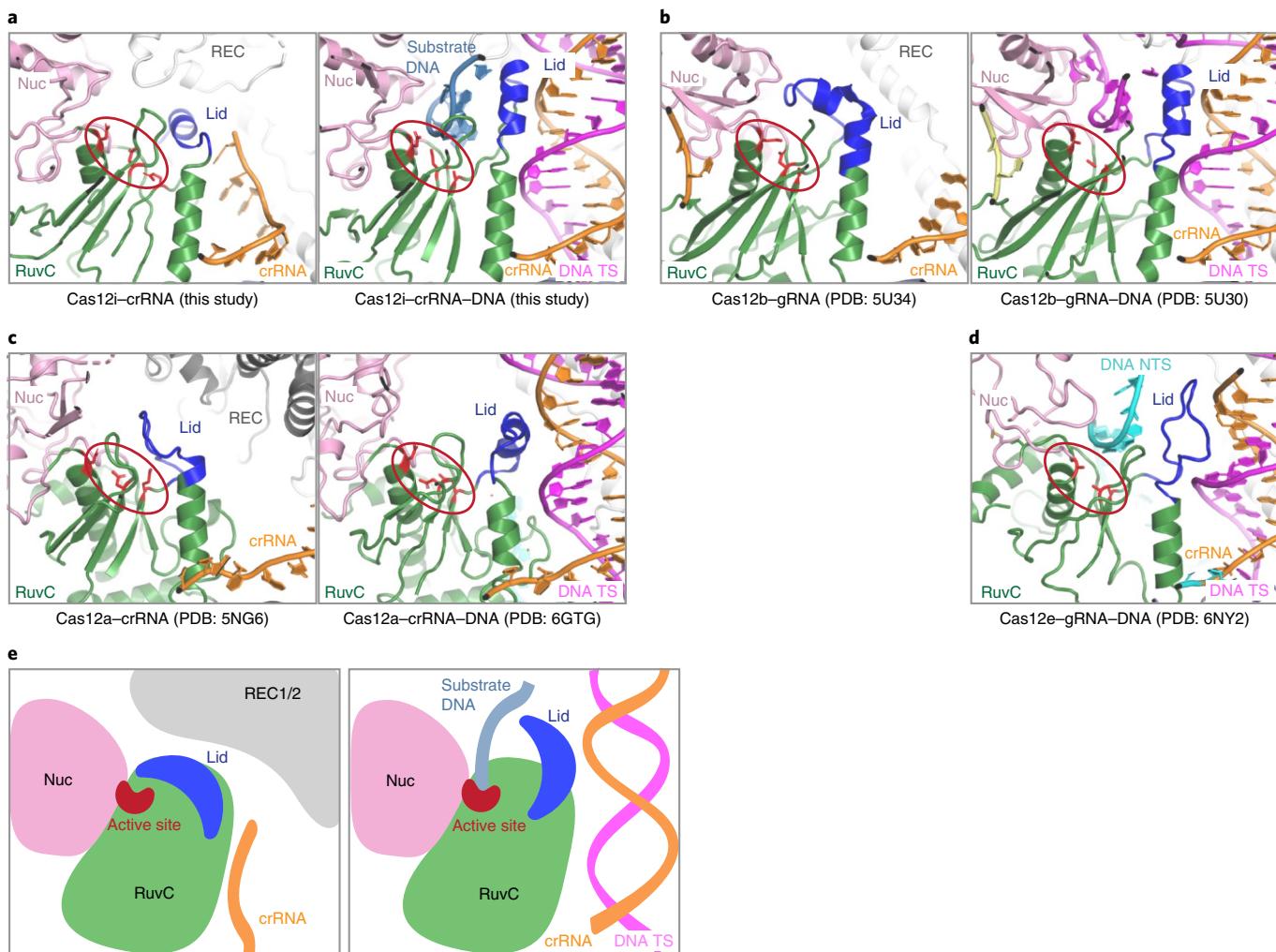


Fig. 4 | The lid in Cas12 endonucleases. **a-d**, The lids (colored in blue) before and after DNA substrate recognition in Cas12i (**a**), Cas12b (**b**), Cas12a (**c**) and Cas12e (**d**) are shown in parallel. Left panels are the structures without substrate (the structure for Cas12e is not available), while right panels are structures with substrate. The RuvC catalytic centers are indicated by red circles, with the three acidic amino acids shown in sticks. **e**, Cartoon model showing that the movement of the lid is coupled with activation of Cas12 endonucleases.

class 2 CRISPR-Cas systems, which not only stabilize the crRNA and allow for recognition by specific CRISPR effectors, but also probably facilitate the correct guide positioning and faithful target binding. Interestingly, six nucleotides immediately adjacent to the spacer-derived guide are identical in Cas12i and Cas12b (Extended Data Fig. 3g). Therefore, the crRNA motif adjacent to the spacer-derived segment may be one signature for classification of CRISPR-Cas systems.

Structures of Cas12i reveal a lid (aa 897–915) located between the last two β -strands (β 4 and β 5) in the RuvC domain and following the second acidic residue of the conserved carboxylate triad, which undergoes a conformational change upon crRNA–DNA heteroduplex formation to facilitate substrate binding and cleavage (Fig. 3e–g). Interestingly, the structural element equivalent to the lid in Cas12i is conserved in Cas12a, Cas12b and Cas12e (Fig. 4a–d), although these Cas12 orthologs have a divergent sequence of the lid region (Extended Data Fig. 8f). First, the lid engages the heteroduplex in a similar sequence-independent manner at positions 8–11. Second, conformational changes of the lid open the RuvC catalytic center. Third, the lid contributes to substrate binding, thereby facilitating DNA cleavage. The lid was demonstrated to be critical for RuvC activation³⁸, suggesting a conserved activation mechanism for Cas12 endonucleases (Fig. 4e).

In summary, our results shed light on the mechanism of the Cas12i system in substrate recognition and activation, and the evolution of the type V endonuclease to acquire distinct functions. Atomic structures would facilitate the manipulation of Cas12i for genome editing with high specificity and diagnostic applications.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41594-020-0499-0>.

Received: 9 March 2020; Accepted: 6 August 2020;

Published online: 07 September 2020

References

1. Mohanraj, P. et al. Diverse evolutionary roots and mechanistic variations of the CRISPR-Cas systems. *Science* **353**, aad5147 (2016).
2. Wright, A. V., Nuñez, J. K. & Doudna, J. A. Biology and applications of CRISPR systems: harnessing Nature’s toolbox for genome engineering. *Cell* **164**, 29–44 (2016).

3. Sorek, R., Lawrence, C. M. & Wiedenheft, B. CRISPR-mediated adaptive immune systems in bacteria and archaea. *Ann. Rev. Biochem.* **82**, 237–266 (2013).
4. Marraffini, L. A. CRISPR-Cas immunity in prokaryotes. *Nature* **526**, 55–61 (2015).
5. Jiang, F. & Doudna, J. A. CRISPR-Cas9 structures and mechanisms. *Annu. Rev. Biophys.* **46**, 505–529 (2017).
6. Hille, F. et al. The biology of CRISPR-Cas: backward and forward. *Cell* **172**, 1239–1259 (2018).
7. Barrangou, R. & Marraffini, L. A. CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity. *Mol. Cell* **54**, 234–244 (2014).
8. van der Oost, J., Jore, M. M., Westra, E. R., Lundgren, M. & Brouns, S. J. CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem. Sci.* **34**, 401–407 (2009).
9. Shmakov, S. et al. Diversity and evolution of class 2 CRISPR-Cas systems. *Nat. Rev. Microbiol.* **15**, 169–182 (2017).
10. Makarova, K. S. et al. An updated evolutionary classification of CRISPR-Cas systems. *Nat. Rev. Microbiol.* **13**, 722–736 (2015).
11. Shmakov, S. et al. Discovery and functional characterization of diverse class 2 CRISPR-Cas systems. *Mol. Cell* **60**, 385–397 (2015).
12. Koonin, E. V., Makarova, K. S. & Zhang, F. Diversity, classification and evolution of CRISPR-Cas systems. *Curr. Opin. Microbiol.* **37**, 67–78 (2017).
13. Barrangou, R. & Doudna, J. A. Applications of CRISPR technologies in research and beyond. *Nat. Biotechnol.* **34**, 933–941 (2016).
14. Komor, A. C., Badran, A. H. & Liu, D. R. CRISPR-based technologies for the manipulation of eukaryotic genomes. *Cell* **168**, 20–36 (2017).
15. Bi, H. et al. CRISPR disruption of *BmOvo* resulted in the failure of emergence and affected the wing and gonad development in the silkworm *Bombyx mori*. *Insects* <https://doi.org/10.3390/insects10080254> (2019).
16. Pickar-Oliver, A. & Gersbach, C. A. The next generation of CRISPR-Cas technologies and applications. *Nat. Rev. Mol. Cell Biol.* **20**, 490–507 (2019).
17. Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).
18. Strecker, J. et al. Engineering of CRISPR-Cas12b for human genome editing. *Nat. Commun.* **10**, 212 (2019).
19. Teng, F. et al. Repurposing CRISPR-Cas12b for mammalian genome engineering. *Cell Discov.* **4**, 63 (2018).
20. Liu, J. J. et al. CasX enzymes comprise a distinct family of RNA-guided genome editors. *Nature* **566**, 218–223 (2019).
21. Yan, W. X. et al. Functionally diverse type V CRISPR-Cas systems. *Science* **363**, 88–91 (2019).
22. Holm, L., Kaariainen, S., Rosenstrom, P. & Schenkel, A. Searching protein structure databases with DaliLite v.3. *Bioinformatics* **24**, 2780–2781 (2008).
23. Swarts, D. C., van der Oost, J. & Jinek, M. Structural basis for guide RNA processing and seed-dependent DNA targeting by CRISPR-Cas12a. *Mol. Cell* **66**, 221–233 (2017).
24. Knott, G. J. et al. Guide-bound structures of an RNA-targeting A-cleaving CRISPR-Cas13a enzyme. *Nat. Struct. Mol. Biol.* **24**, 825–833 (2017).
25. 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).
26. Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K. & Doudna, J. A. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **329**, 1355–1358 (2010).
27. Liu, L. et al. The molecular architecture for RNA-guided RNA cleavage by Cas13a. *Cell* **170**, 714–726 (2017).
28. Yang, H., Gao, P., Rajashankar, K. R. & Patel, D. J. PAM-dependent target DNA recognition and cleavage by C2c1 CRISPR-Cas endonuclease. *Cell* **167**, 1814–1828 (2016).
29. Liu, L. et al. C2c1-sgRNA complex structure reveals RNA-guided DNA cleavage mechanism. *Mol. Cell* **65**, 310–322 (2017).
30. Swarts, D. C. Stirring up the type V alphabet soup. *CRISPR J.* **2**, 14–16 (2019).
31. Strecker, J. et al. RNA-guided DNA insertion with CRISPR-associated transposases. *Science* **365**, 48–53 (2019).
32. Al-Shayeb, B. et al. Clades of huge phages from across Earth's ecosystems. *Nature* **578**, 425–431 (2020).
33. Dong, D. et al. The crystal structure of Cpf1 in complex with CRISPR RNA. *Nature* **532**, 522–526 (2016).
34. Gao, P., Yang, H., Rajashankar, K. R., Huang, Z. & Patel, D. J. Type V CRISPR-Cas Cpf1 endonuclease employs a unique mechanism for crRNA-mediated target DNA recognition. *Cell Res.* **26**, 901–913 (2016).
35. Yamano, T. et al. Crystal structure of Cpf1 in complex with guide RNA and target DNA. *Cell* **165**, 949–962 (2016).
36. Yamano, T. et al. Structural basis for the canonical and non-canonical PAM recognition by CRISPR-Cpf1. *Mol. Cell* **67**, 633–645 (2017).
37. Stella, S., Alcón, P. & Montoya, G. Structure of the Cpf1 endonuclease R-loop complex after target DNA cleavage. *Nature* **546**, 559–563 (2017).
38. Stella, S. et al. Conformational activation promotes CRISPR-Cas12a catalysis and resetting of the endonuclease activity. *Cell* **175**, 1856–1871 (2018).
39. Nishimasu, H. et al. Structural basis for the altered PAM recognition by engineered CRISPR-Cpf1. *Mol. Cell* **67**, 139–147 (2017).
40. Swarts, D. C. & Jinek, M. Mechanistic Insights into the *cis*- and *trans*-acting DNase activities of Cas12a. *Mol. Cell* **73**, 589–600 (2019).
41. Zhang, H. et al. Structural basis for the inhibition of CRISPR-Cas12a by anti-CRISPR proteins. *Cell Host Microbe* **25**, 815–826 (2019).
42. Wu, D., Guan, X., Zhu, Y., Ren, K. & Huang, Z. Structural basis of stringent PAM recognition by CRISPR-C2c1 in complex with sgRNA. *Cell Res.* **27**, 705–708 (2017).
43. Nishimasu, H. et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **156**, 935–949 (2014).
44. Anders, C., Niewoehner, O., Duerst, A. & Jinek, M. Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513**, 569–573 (2014).
45. Ran, F. A. et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **154**, 1380–1389 (2013).
46. Kulcsár, P. I. et al. Blackjack mutations improve the on-target activities of increased fidelity variants of SpCas9 with 5'G-extended sgRNAs. *Nat. Commun.* <https://doi.org/10.1038/s41467-020-15021-5> (2020).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2020

Methods

Protein expression and purification. The plasmid encoding full-length Cas12i (Cas12i) was obtained from Addgene (120882)²¹. Cas12i was fused with an N-terminal 6xHis tag. The mutations were generated by quick change mutagenesis. These plasmids were transformed into BL21-CodonPlus (DE3)-RIL cells for overexpression in Terrific Broth medium. Protein expression was induced by adding 0.5 mM IPTG at 18 °C overnight. The cells were collected and then disrupted by sonication in buffer A (25 mM Tris-HCl, pH 8.0, 500 mM NaCl, 5% glycerol, 1 mM PMSF and 5 mM β-mercaptoethanol). After centrifugation, the supernatants were incubated with the Ni-NTA resin and washed with buffer B (25 mM Tris-HCl, pH 8.0, 500 mM NaCl, 30 mM imidazole and 5 mM β-mercaptoethanol), then the His-tagged target proteins were eluted using buffer B supplemented with 300 mM imidazole. Subsequently, the target proteins were loaded onto a heparin column (GE) and eluted with a linear sodium chloride gradient, followed by size exclusion chromatography (Superdex 200, GE) in buffer C (25 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2 mM DTT and 0.1 mM MgCl₂). Peak fractions were concentrated and stocked in -80 °C.

To assemble the Cas12i-crRNA binary complex, Cas12i proteins were mixed with crRNA (synthesized from Integrated DNA Technologies) at a ratio of 1:1.3 on ice for 1 h. To reconstitute the Cas12i-crRNA-target DNA ternary complexes, the Cas12i WT or E894A mutant proteins were incubated with crRNA for 30 min, followed by adding the target DNA synthesized from IDT at a ratio of 1:1.3:2. The mixture was applied to a Superdex 200 (GE) column equilibrated with buffer C.

In vitro DNA cleavage assay. Target DNA containing the 5'-TTN-3' PAM was synthesized from IDT and cloned into pET28-MHL vector using the InFusion cloning kit (ClonTech). Poly-A, -T, -C and -G substrates were generated by substitution at the TS of the protospacer-derived sequence at positions 17–28. Cas12i proteins (250 ng) were mixed with crRNA at a ratio of 1:1.5 on ice for 30 min, in cleavage buffer containing 25 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2 mM MgCl₂, 5 mM DTT and 5% glycerol. The purified plasmids (100 ng) were then added into the mixture for another 30 min. The reactions were quenched by adding EDTA and proteinase K. The cleavage products were resolved on 0.5% agarose gels and visualized by ethidium bromide staining.

Pre-crRNA cleavage assays. The 60-nt pre-crRNA synthesized from IDT was incubated with 2 µg of Cas12i proteins on ice at a ratio of 1:3 in cleavage buffer. The reactions were stopped by adding Tris-borate-EDTA(TBE)-urea sample buffer (Thermo Fisher Scientific) at 66 °C and quenching for 10 min. The products were separated on a 15% TBE-urea gel and stained using SYBR Gold (Invitrogen).

Electron microscopy. Aliquots of 3-µl samples (Cas12i-crRNA, Cas12i^{WT}-crRNA-dsDNA and Cas12i^{E894A}-crRNA-dsDNA) at 0.75 mg ml⁻¹ were applied to glow-discharged Quantifoil holey carbon girds (Au, R1.2/1.3, 300 mesh). The grids were blotted for 4 s and plunged into liquid ethane using a Gatan Cryoplunge 3 plunger. Cryo-EM data were collected with a Titan Krios microscope (FEI) operated at 300 kV and images were collected using Leginon⁴⁷ at a nominal magnification of ×81,000 (resulting in a calibrated physical pixel size of 1.05 Å per pixel) with a defocus range of 1.2–2.5 µm. The images were recorded on a K2 summit (for wild-type Cas12i-crRNA-dsDNA) or K3 (for Cas12i^{E894A}-crRNA-dsDNA) electron direct detector in super-resolution mode at the end of a GIF-Quantum energy filter operated with a slit width of 20 eV. For data collection with a K2 detector, a dose rate of 5 electrons per pixel per second and an exposure time of 8 s were used, generating 40 video frames (200 ms per frame) with a total dose of ~35 electrons per Å². For data collection with a K3 detector, a dose rate of 20 electrons per pixel per second and an exposure time of 3.12 s were used, generating 40 video frames (78 ms per frame) with a total dose of ~54 electrons per Å². The statistics for the cryo-EM data are listed in Table 1.

Image processing. The video frames were imported to RELION-3⁴⁸. Video frames were aligned using MotionCor2⁴⁹ with a binning factor of 2. Contrast transfer function (CTF) parameters were estimated using Gctf⁵⁰. A few thousand particles were auto-picked without template to generate 2D averages for subsequent template-based auto-picking.

For the Cas12i^{E894A}-crRNA-dsDNA dataset, five million particles were auto-picked and extracted from the dose weighted micrographs. The dataset was split into batches for 2D classifications, which were used to exclude false and bad particles that fall into 2D averages with poor features. Particles from different views were used to generate the initial model in cryoSPARC⁵¹. The dataset was split into six batches for 3D refinement, each converging at a resolution of around 3.3 Å. After Bayesian polishing, each part was classified into six classes in 3D classification. The 3D classification shows that the complex adopts a single conformation. The classes representing bad particles from Bayesian polishing and the classes with less alignment accuracy were excluded. The remaining 731,231 particles were used for 3D refinement, CTF refinement and 3D refinement, which finally converged at 2.9-Å resolution. Focused 3D refinement, focused 3D classification and 3D refinement were performed sequentially around the PAM binding region to improve the local map quality.

For the Cas12i^{WT}-crRNA-dsDNA dataset, 1.5 million particles were auto-picked and extracted from the dose weighted micrographs. The dataset was split into batches for 2D classifications, which were used to exclude false and bad particles that fall into 2D averages with poor features. Particles from different views were used to generate the initial model in cryoSPARC⁵¹. The dataset was split into two batches for 3D refinement, each converging at a resolution of around 4 Å. After Bayesian polishing, each part was classified into six classes in 3D classification. The 3D classification shows that the complex adopts a different conformation. Cas12i accommodates the crRNA-DNA duplex with different lengths, and different conformational states were distinguished accordingly. The dataset representing the Cas12i-crRNA binary complex contained 122,435 particles, converging at a resolution of 3.9 Å. The dataset representing the I1 state (14–15 bp crRNA-DNA heteroduplex) contained 75,406 particles, converging at a resolution of 3.9 Å. The dataset representing the Cas12i-crRNA-DNA ternary complex (26–28 bp crRNA-DNA heteroduplex) contained 99,516 particles, resulting in a 3.8-Å resolution map similar to the map of Cas12i^{E894A}-crRNA-dsDNA.

The 3D FSC analyses for the presented maps were performed using the Remote 3DFSC Processing Server (<https://3dfsc.salk.edu/upload/>)⁵². Map-to-model FSC analyses were performed using the phenix.validation_cryoem tool in Phenix.

Model building and refinement. De novo model building of the Cas12i^{E894A}-crRNA-dsDNA structure was performed manually in COOT⁵³. Secondary structure predictions by PSIPRED⁵⁴ were used to assist manual building. Refinement of the structure models against corresponding maps was performed using the phenix.real_space_refine tool in Phenix⁵⁵. For the Cas12i-crRNA binary complex and the I1 state, the Cas12i^{E894A}-crRNA-dsDNA structure was first docked to the corresponding cryo-EM maps using UCSF Chimera⁵⁶, then manually adjusted in COOT, and finally refined in Phenix.

Visualization. Figures were generated using PyMOL and UCSF Chimera⁵⁶.

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

Data availability

Cryo-EM reconstructions of Cas12i(E894A)-crRNA-dsDNA, Cas12i-crRNA and the I1 complexes have been deposited in the Electron Microscopy Data Bank under accession nos. EMD-21541, EMD-21551 and EMD-21552, respectively. Coordinates for atomic models of Cas12i(E894A)-crRNA-dsDNA, Cas12i-crRNA and the I1 complexes have been deposited in the Protein Data Bank under accession nos. 6W5C, 6W62 and 6W64, respectively. Source data are provided with this paper.

References

47. Suloway, C. et al. Automated molecular microscopy: the new Leginon system. *J. Struct. Biol.* **151**, 41–60 (2005).
48. Zivanov, J. et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. *eLife* <https://doi.org/10.7554/eLife.42166> (2018).
49. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* **14**, 331–332 (2017).
50. Zhang, K. Gctf: real-time CTF determination and correction. *J. Struct. Biol.* **193**, 1–12 (2016).
51. 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).
52. Tan, Y. Z. et al. Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods* **14**, 793–796 (2017).
53. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).
54. Jones, D. T. Protein secondary structure prediction based on position-specific scoring matrices. *J. Mol. Biol.* **292**, 195–202 (1999).
55. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr. D Struct. Biol.* **74**, 531–544 (2018).
56. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

Acknowledgements

We thank T. Klose and V. Bowman for help with cryo-EM, S. Wilson for computation and J. Tesmer and C. Gabel for critical reading of the manuscript. This work was supported by NIH grant no. R01GM138675 and a Showalter Trust Research Award to L.C.

Author contributions

H.Z., R.X. and Z.L. prepared samples. Z.L., H.Z. and L.C. collected and processed cryo-EM data. H.Z. and R.X. performed biochemical analysis. Z.L., H.Z. and L.C.

prepared figures. All authors analyzed the data. H.Z., Z.L. and L.C. prepared the manuscript with input from R.X.

Competing interests

The authors declare no competing interests.

Additional information

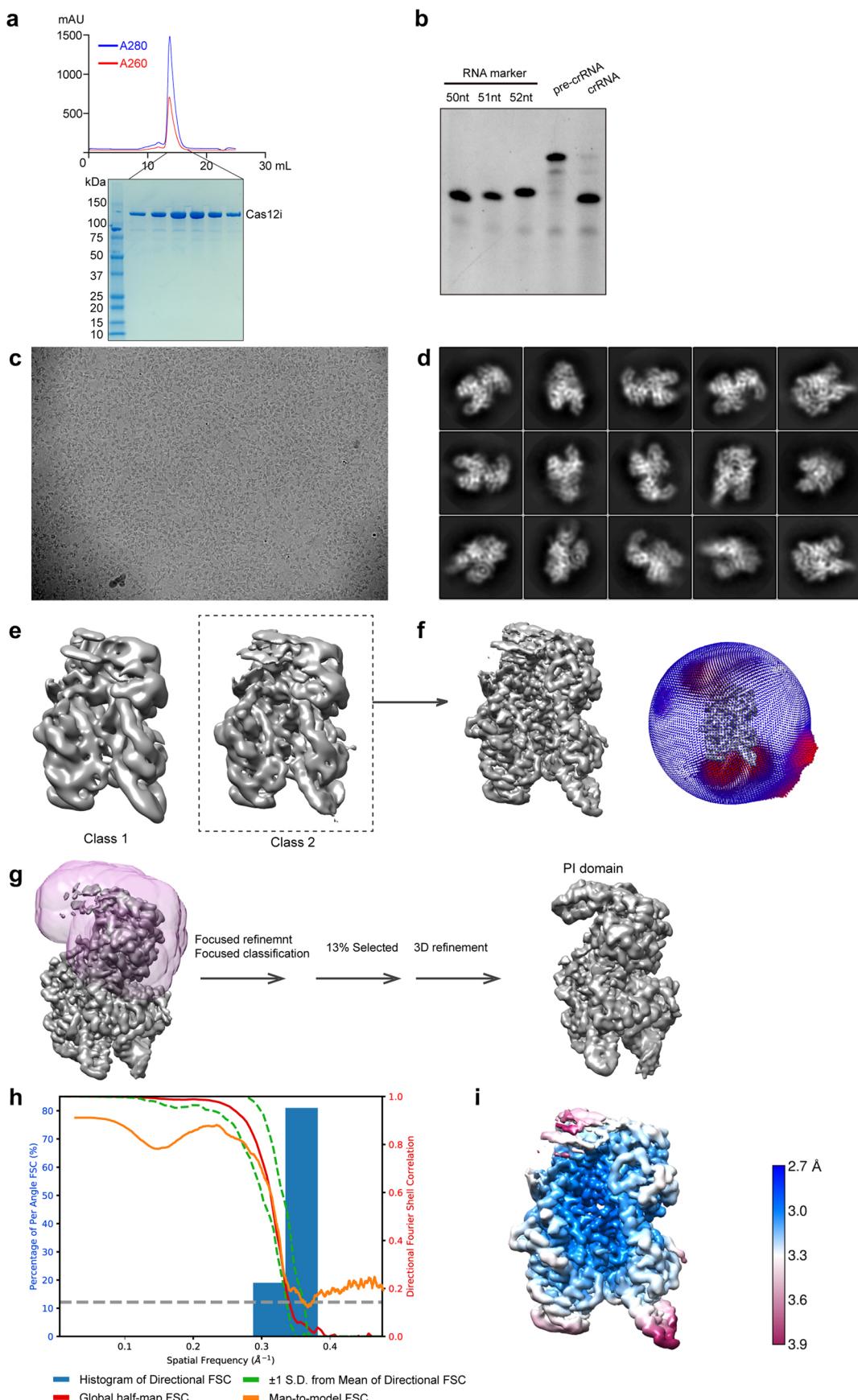
Extended data is available for this paper at <https://doi.org/10.1038/s41594-020-0499-0>.

Supplementary information is available for this paper at <https://doi.org/10.1038/s41594-020-0499-0>.

Correspondence and requests for materials should be addressed to H.Z. or L.C.

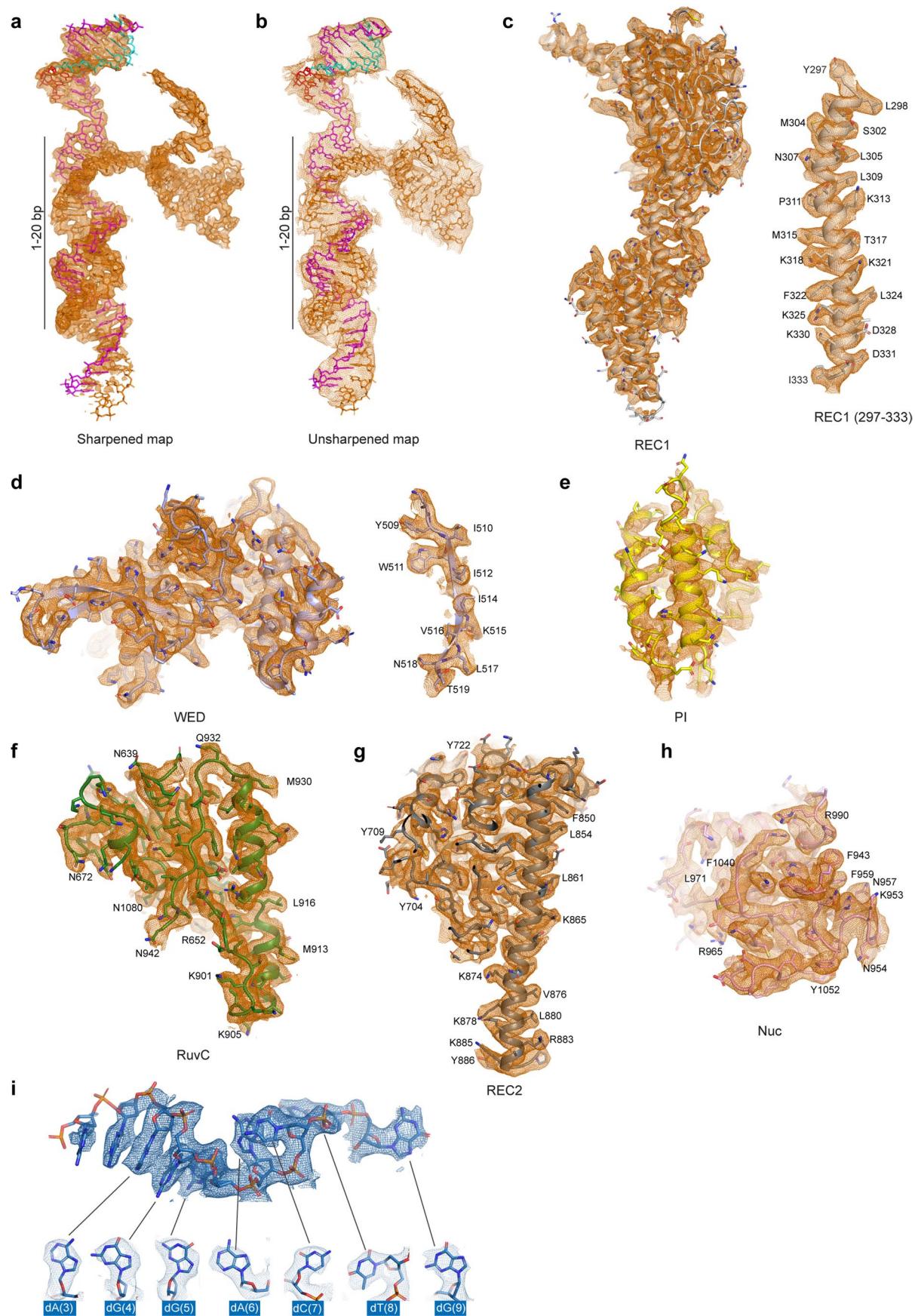
Peer review information Katarzyna Marcinkiewicz and Anke Sparmann were the primary editors on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.



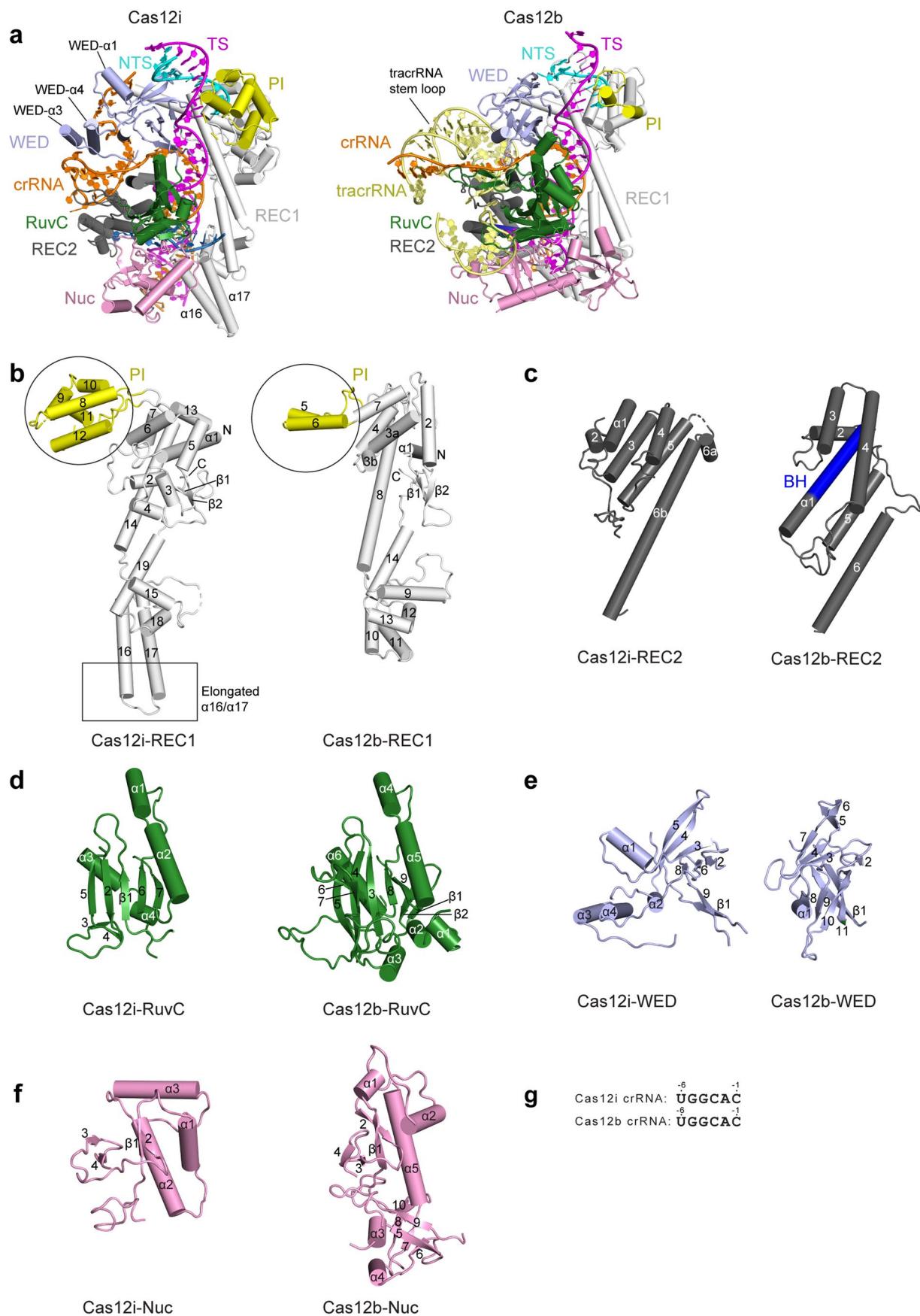
Extended Data Fig. 1 | See next page for caption.

Extended Data Fig. 1 | Sample preparation and cryo-EM for the Cas12i(E894A)-crRNA-dsDNA complex. **a**, Purification of Cas12i. Upper: Size exclusion chromatography (SEC) profile of Cas12i. UV absorbance curves at 280 nm and 260 nm are shown in blue and red, respectively. Lower: SDS-PAGE analysis of the elution fractions from SEC as indicated. **b**, Urea-PAGE analysis of pre-crRNA before and after processing by Cas12i. RNA markers indicate that the mature crRNA is 51nt. **c**, A representative raw cryo-EM micrograph of the Cas12i-crRNA-dsDNA ternary complex. **d**, Representative 2D class averages. **e**, Two major classes from 3D classification. The two maps are similar, with class 2 at better quality. **f**, 3D auto-refinement using particles from class 2. Angular distribution of particles is shown on the right. **g**, Procedures used for improving the resolution of the PI domain. Focused refinement followed by focused 3D classification using a soft mask (in pink mesh) resulted in a 3D class from 13% particles with improved density in the PI domain. **h**, Plot of the global half-map FSC (solid red line), map-to-model FSC (solid orange line), and spread of directional resolution values ($\pm 1\sigma$ from mean, green dotted lines); the blue bars indicate a histogram of 100 such values evenly sampled over the 3D FSC. **i**, Local resolution map of the final reconstruction. Uncropped images for panels a and b are available as source data online.



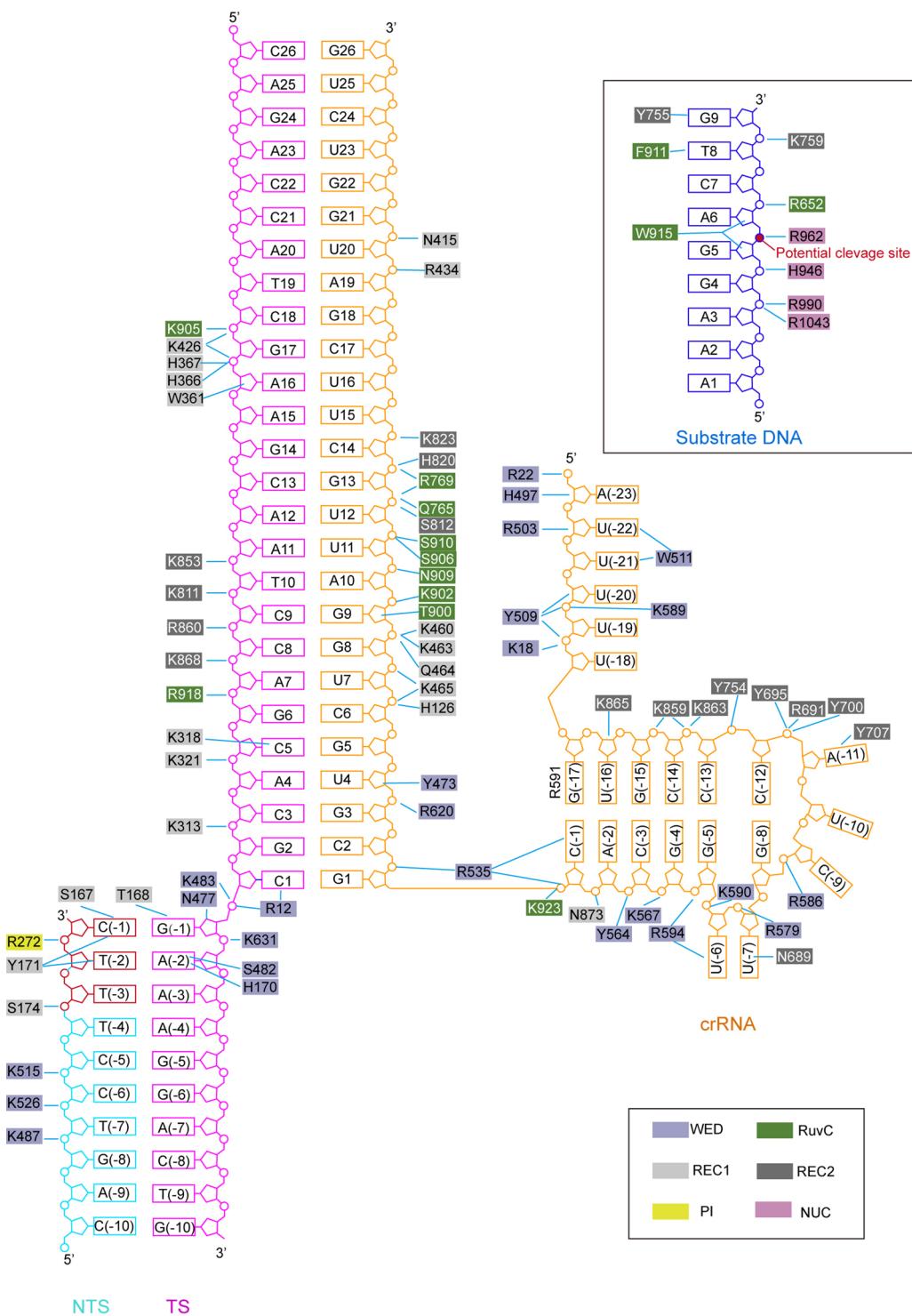
Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | Detailed cryo-EM density map of the Cas12i(E894A)-crRNA-dsDNA complex with final atomic model fitted in. **a, b**, Fitting of nucleic acids to the corresponding cryo-EM map. The atomic models are shown in stick with crRNA, the target strand and the non-target strand colored in orange, magenta and cyan, respectively. Cryo-EM density from the sharpened map (**a**) or the unsharpened map (**b**) is shown in mesh. **c**, Fitting of the REC1 domain. A representative α -helix from the REC1 domain is shown in details on the right. **d**, Fitting of the WED domain. A representative β -strand is shown in details on the right. **e**, Fitting of the PI domain. **f**, Fitting of the RuvC domain. **g**, Fitting of the REC2 domain. **h**, Fitting of the Nuc domain. **i**, Fitting of the substrate DNA.

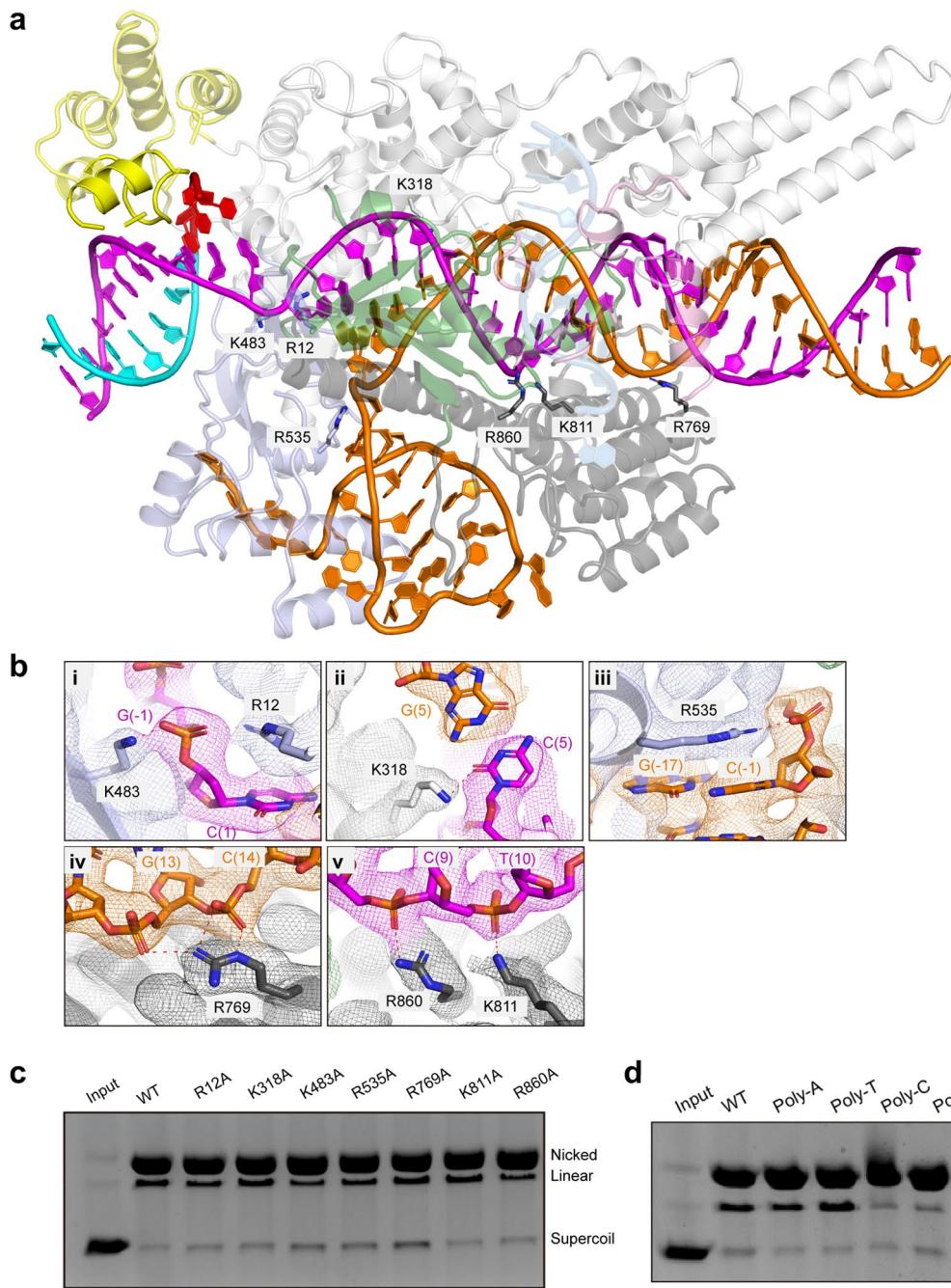


Extended Data Fig. 3 | See next page for caption.

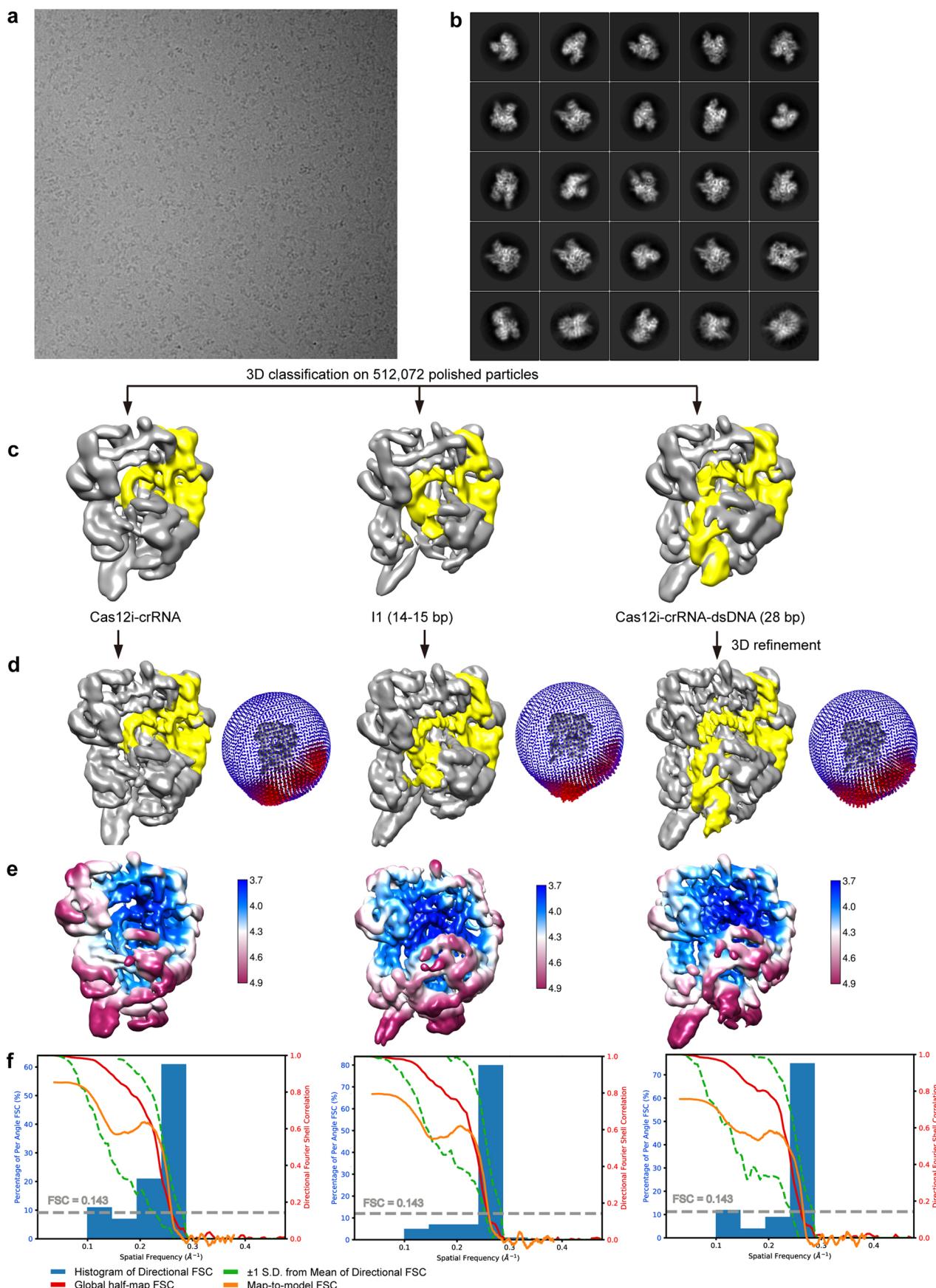
Extended Data Fig. 3 | Structural comparison of Cas12i and Cas12b. **a**, Overall structures of the Cas12i-crRNA-dsDNA and the Cas12b-gRNA-dsDNA complexes. The structures are aligned by secondary-structure matching (SSM) in COOT. **b-f**, Structural comparison of each domain. Secondary structures in each domain are labeled. The PI domain and the extended α 16 and α 17 helix pair are indicated by circles and a box, respectively. **g**, Sequence alignment of the crRNA repeat adjacent to the spacer-derived segment between Cas12i and Cas12b.



Extended Data Fig. 4 | Schematic of nucleic acid recognition in the Cas12i(E894A)-crRNA-dsDNA complex. Intermolecular contacts between Cas12i and nucleic acids, including the target and non-target strands and the substrate DNA, are shown by solid lines. The Cas12i residues are colored based on the domain architecture.

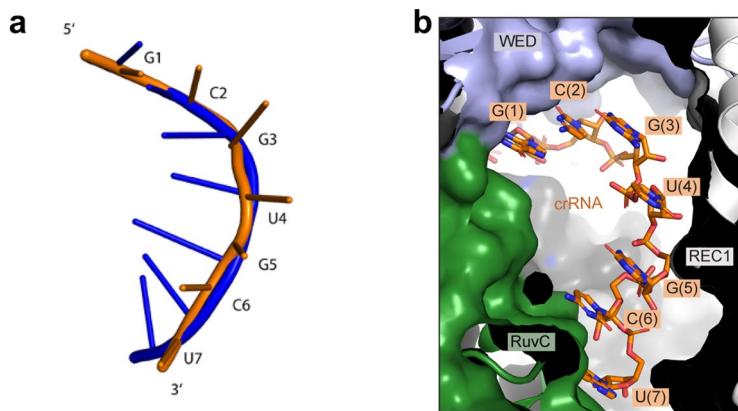


Extended Data Fig. 5 | Structure and mutagenesis analysis for the recognition of the crRNA-target DNA heteroduplex. **a**, Structure of the Cas12i-crRNA-DNA complex with selected key residues involved in the recognition of the heteroduplex shown as sticks. **b**, Close-up view of the interactions shown in a with cryo-EM map shown in mesh. **c**, Substrate cleavage assay using wild-type Cas12i and Cas12i with single mutations on the key residues shown in a. The results shown are representative of three experiments. **d**, Substrate cleavage assay using poly-A, -T, -C, and G as substrates. The results shown are representative of three experiments. Uncropped images for panels c and d are available as source data online.

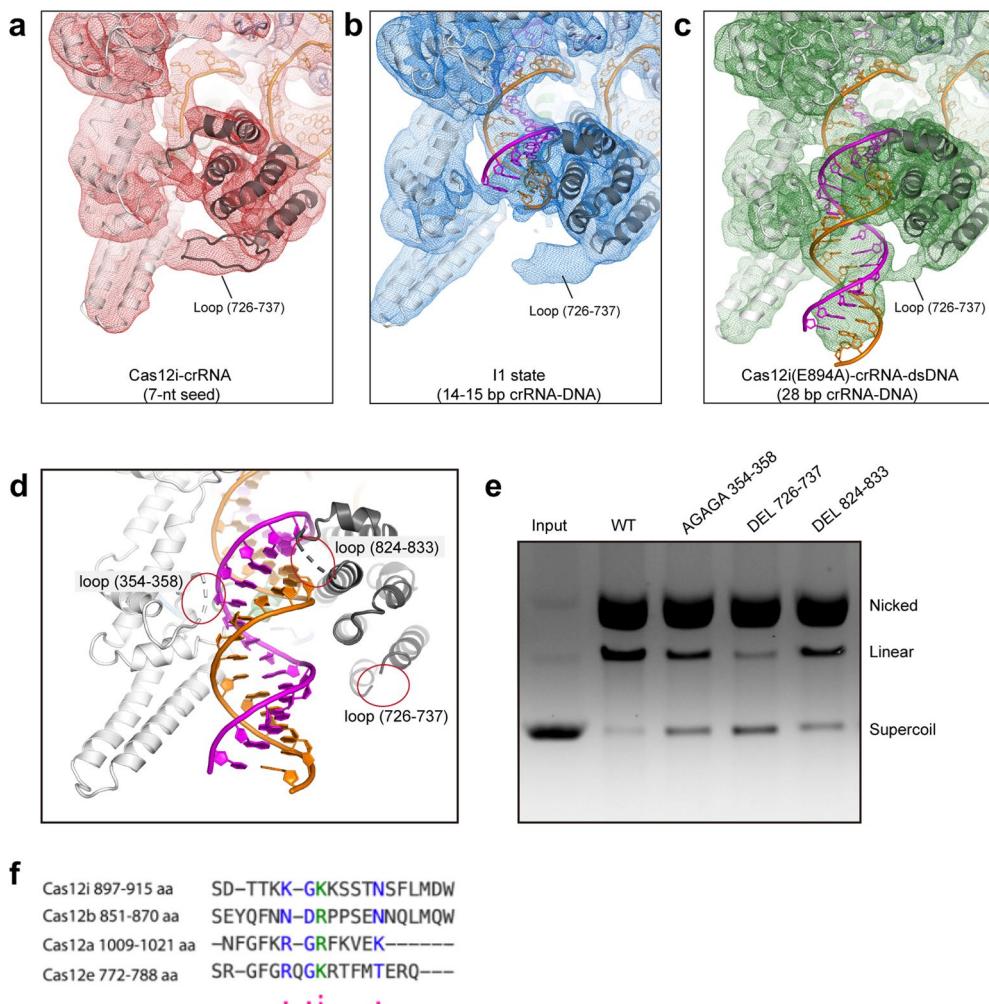


Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Cryo-EM data processing for the wild-type Cas12i-crRNA-dsDNA complex. **a**, A representative raw cryo-EM micrograph of the wild-type Cas12i-crRNA-DNA complex. **b**, Representative 2D class averages. **c**, 3D classification. Three major classes are observed, representing the Cas12i-crRNA binary complex, the intermediate state (I1 state, 14–15 bp heteroduplex), and the fully assembled Cas12i-crRNA-DNA complex (26–28 bp heteroduplex). Densities corresponding to the crRNA and target DNA are colored in yellow. The Cas12i-crRNA map is comparable to a reconstruction using biochemically purified Cas12i-crRNA complex (data not shown). **d**, 3D auto-refinement for the three classes shown in c. Angular distribution of each reconstruction is shown below. **e**, Local resolution map of the final reconstructions in d. **f**, Plot of the global half-map FSC (solid red line), map-to-model FSC (solid orange line), and spread of directional resolution values ($\pm 1\sigma$ from mean, green dotted lines; the blue bars indicate a histogram of 100 such values evenly sampled over the 3D FSC).



Extended Data Fig. 7 | Structural analysis of the seed sequence. **a**, Superposition of the seed sequence of Cas12i (orange) with the same sequence simulated in an A-form geometry (blue). **b**, Structure of the seed sequence with Cas12i domains shown in surface representation.



Extended Data Fig. 8 | Conformational changes of recognition loops in Cas12i upon heteroduplex formation. **a**, Structure of the Cas12i-crRNA binary complex with cryo-EM density shown in mesh. Loop 726-737 is indicated. **b**, Structure of the I1 state with cryo-EM density shown in mesh. **c**, Structure of the Cas12i(E894A)-crRNA-DNA ternary complex with cryo-EM density shown in mesh. **d**, Structure showing three loops from the REC1 and REC2 domains of Cas12i that are likely involved in heteroduplex recognition. **e**, Substrate cleavage assay using wild-type Cas12i and Cas12i with each of the three loops shown in **d** deleted or mutated. The results shown are representative of three experiments. **f**, Sequence alignment of the lid region among Cas12 orthologs using ClustalW program (Thompson, J. D., Gibson, T. J. & Higgins, D. G. Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics Chapter 2*, Unit 2 3, doi:10.1002/0471250953.bi0203s00 (2002).). Uncropped image for panel **e** is available as source data online.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Legion (J Struct Biol 151, 41-60)

Data analysis

Appion 3.3, Relion 3.0, Gctf 1.06, Motioncor2 1.1.0, PyMOL 2.2.3, Chimera 1.13, Coot 0.8.9.1, Phenix 1.12

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All cryo-EM maps and coordinates have been deposited to EMDB and PDB.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No sample-size calculation was performed. The actual sample size was increased so that desired structure features were obtained in cryo-EM maps.
Data exclusions	No data were excluded.
Replication	All attempts of replication were successful.
Randomization	Not relevant to this study.
Blinding	Not relevant to this study.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems		Methods	
n/a	Involved in the study	n/a	Involved in the study
<input checked="" type="checkbox"/>	Antibodies	<input checked="" type="checkbox"/>	ChIP-seq
<input checked="" type="checkbox"/>	Eukaryotic cell lines	<input checked="" type="checkbox"/>	Flow cytometry
<input checked="" type="checkbox"/>	Palaeontology	<input checked="" type="checkbox"/>	MRI-based neuroimaging
<input checked="" type="checkbox"/>	Animals and other organisms		
<input checked="" type="checkbox"/>	Human research participants		
<input checked="" type="checkbox"/>	Clinical data		