structural comparison of Cascade to eukaryotic Argonautes reveals a similar "kink helix" positioned between nucleotides 6 and 7 (fig. S13) (*36*). However, in Argonautes there is no thumb that covers the kinked base, and it is expected that target hybridization may release the RNA for contiguous duplex formation (*37*). Recent crystal structures of the Cas9 protein suggest a similar protein-mediated preordering of the RNA guide (*38*, *39*), and a structure of the target-bound complex suggests that the RNA-DNA hybrid forms a contiguous A-form duplex (*39*).

Target detection by Cascade relies on protein-mediated recognition of a three-nucleotide PAM and crRNA-guided hybridization to the target. PAM recognition has been proposed to destabilize the target DNA duplex and initiate crRNA-guided strand invasion. L1 in Cse1 has been implicated in this process, and the structure explains why mutations in L1 result in Cascade assembly defects (Fig. 5) (*31*). However, the structure of Cascade without DNA does not explain how Cascade recognizes the PAM. Structures of Cascade, in association with DNA and Cas3, may provide additional insights into the interplay of Cascade and Cas3 in the process of RNA-guided DNA interference.

REFERENCES AND NOTES

1. R. Sorek, C. M. Lawrence, B. Wiedenheft, *Annu. Rev. Biochem.* **82**, 237–266 (2013).
2. J. van der Oost, E. R. Westra, R. N. Jackson, B. Wiedenheft, *Nat. Rev. Microbiol.* **12**, 479–492 (2014).
3. J. Bondy-Denomy, A. R. Davidson, *Trends Microbiol.* **22**, 218–225 (2014).
4. B. Wiedenheft, S. H. Sternberg, J. A. Doudna, *Nature* **482**, 331–338 (2012).
5. J. Reeks, J. H. Naismith, M. F. White, *Biochem. J.* **453**, 155–166 (2013).
6. K. S. Makarova *et al.*, *Nat. Rev. Microbiol.* **9**, 467–477 (2011).
7. S. J. Brouns *et al.*, *Science* **321**, 960–964 (2008).
8. M. M. Jore *et al.*, *Nat. Struct. Mol. Biol.* **18**, 529–536 (2011).
9. E. R. Westra *et al.*, *PLOS Genet.* **9**, e1003742 (2013).
10. E. Semenova *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10098–10103 (2011).
11. M. L. Hochstrasser *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 6618–6623 (2014).
12. B. Wiedenheft *et al.*, *Nature* **477**, 486–489 (2011).
13. E. R. Westra *et al.*, *Mol. Cell* **46**, 595–605 (2012).
14. R. N. Jackson, M. Lavin, J. Carter, B. Wiedenheft, *Curr. Opin. Struct. Biol.* **24**, 106–114 (2014).
15. E. R. Westra *et al.*, *Annu. Rev. Genet.* **46**, 311–339 (2012).
16. T. C. Terwilliger, *J. Struct. Funct. Genomics* **14**, 91–95 (2013).
17. E. M. Gesner, M. J. Schellenberg, E. L. Garside, M. M. George, A. M. Macmillan, *Nat. Struct. Mol. Biol.* **18**, 688–692 (2011).
18. D. G. Sashital, M. Jinek, J. A. Doudna, *Nat. Struct. Mol. Biol.* **18**, 680–687 (2011).
19. J. Carte, R. Wang, H. Li, R. M. Terns, M. P. Terns, *Genes Dev.* **22**, 3489–3496 (2008).
20. R. E. Haurwitz, M. Jinek, B. Wiedenheft, K. Zhou, J. A. Doudna, *Science* **329**, 1355–1358 (2010).
21. A. Ebihara *et al.*, *Protein Sci.* **15**, 1494–1499 (2006).
22. Single-letter abbreviations for the amino acid residues are as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; and Y, Tyr.
23. A. Hrle *et al.*, *RNA Biol.* **10**, 1670–1678 (2013).
24. N. G. Lintner *et al.*, *J. Biol. Chem.* **286**, 21643–21656 (2011).
25. B. Wiedenheft *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10092–10097 (2011).
26. T. Künne, D. C. Swarts, S. J. Brouns, *Trends Microbiol.* **22**, 74–83 (2014).
27. K. H. Nam, Q. Huang, A. Ke, *FEBS Lett.* **586**, 3956–3961 (2012).
28. E. L. Garside *et al.*, *RNA* **18**, 2020–2028 (2012).
29. Y. Koo, D. Ka, E. J. Kim, N. Suh, E. Bae, *J. Mol. Biol.* **425**, 3799–3810 (2013).
30. K. H. Nam *et al.*, *Structure* **20**, 1574–1584 (2012).
31. D. G. Sashital, B. Wiedenheft, J. A. Doudna, *Mol. Cell* **46**, 606–615 (2012).
32. S. Mulepati, A. Orr, S. Bailey, *J. Biol. Chem.* **287**, 22445–22449 (2012).
33. C. Rouillon *et al.*, *Mol. Cell* **52**, 124–134 (2013).
34. M. Spilman *et al.*, *Mol. Cell* **52**, 146–152 (2013).
35. R. H. Staals *et al.*, *Mol. Cell* **52**, 135–145 (2013).
36. N. T. Schirle, I. J. MacRae, *Science* **336**, 1037–1040 (2012).
37. G. Sheng *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 652–657 (2014).
38. M. Jinek *et al.*, *Science* **343**, 1247997 (2014).
39. H. Nishimasu *et al.*, *Cell* **156**, 935–949 (2014).

ACKNOWLEDGMENTS

We thank J. Richardson and D. Richardson for technical suggestions and discussion and A. McCoy for implementing the EM scale factor refinement in Phaser. X-ray diffraction data was collected with assistance from J. Nix at Advanced Light Source (ALS) beamline 4.2.2 (DE-AC02-05CH11231), R. Sanishvili and C. Ogata at Advanced Photon Source (APS) beamline 23-ID (Y1-GM-1104), the Structural Biology Center at APS 19-ID (DE-AC02-06CH11357), and Stanford Synchrotron Radiation Lightsource (DE-AC02-76SF00515 and P41GM103393). E.R.W. received funding from the People Program (Marie Curie Actions) of the European Union's Seventh Framework Program (FP7/2007-2013) under REA grant agreement 327606. S.J.J. B. is supported by a Vidi grant from the Netherlands Organization of Scientific Research (864.11.005) and J.v.d.O. by a Vici grant (865.05.001). R.J.R. is supported by a Principal Research Fellowship from the Wellcome Trust (grant 082961/Z/07/Z). T.C.T. and R.J.R. are supported by a grant (GM063210) from the NIH. J.C. is supported by a grant for undergraduate research from the Howard Hughes Medical Institute (52006931). R.N.J. is supported by a National Research Service Award postdoctoral fellowship (F32 GM108436) from the NIH. Research in the Wiedenheft lab is supported by the NIH (P20GM103500 and R01GM108888), the NSF Experimental Program to Stimulate Competitive Research (EPS-110134), the M. J. Murdock Charitable Trust, and the Montana State University Agricultural Experiment Station. Atomic coordinates have been deposited into the Protein Data Bank with accession code 4TVX.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6203/1473/suppl/DC1
Materials and Methods
Figs. S1 to S13
Tables S1 and S2
References (*40*–*52*)
Movie S1

21 May 2014; accepted 24 July 2014
Published online 7 August 2014;
10.1126/science.1256328

STRUCTURAL BIOLOGY

Crystal structure of a CRISPR RNA–guided surveillance complex bound to a ssDNA target

Sabin Mulepati,[1] Annie Héroux,[2] Scott Bailey[1]*

In prokaryotes, RNA derived from type I and type III CRISPR loci direct large ribonucleoprotein complexes to destroy invading bacteriophage and plasmids. In *Escherichia coli*, this 405-kilodalton complex is called Cascade. We report the crystal structure of Cascade bound to a single-stranded DNA (ssDNA) target at a resolution of 3.03 angstroms. The structure reveals that the CRISPR RNA and target strands do not form a double helix but instead adopt an underwound ribbon-like structure. This noncanonical structure is facilitated by rotation of every sixth nucleotide out of the RNA-DNA hybrid and is stabilized by the highly interlocked organization of protein subunits. These studies provide insight into both the assembly and the activity of this complex and suggest a mechanism to enforce fidelity of target binding.

Prokaryotes use an RNA-based adaptive immune system called the CRISPR (clustered regularly interspaced short palindromic repeats)–Cas (CRISPR-associated) system to prevent invasion by bacteriophage and plasmids (*1*, *2*). Found in the host genome, CRISPR loci comprise an array of identical repeats interspersed by variable foreign DNA. CRISPR loci are transcribed and then cleaved within the repeat regions to generate small CRISPR RNAs (crRNAs) (*3*–*5*). In type I and type III CRISPR-Cas systems (*6*), crRNA and Cas proteins assemble into large multisubunit complexes (*3*, *7*–*12*). Despite a common seahorse architecture (*8*, *13*–*17*), these complexes use different mechanisms to destroy their targets: Type I complexes are surveillance complexes that bind DNA complementary to the crRNA guide and then recruit a trans-acting helicase-nuclease, Cas3, to unwind and degrade the invading DNA (*18*–*20*), whereas type III-A complexes are effector complexes that destroy targets directly via their intrinsic nuclease activity (*7*, *10*).

The type I-E surveillance complex in *E. coli* is known as Cascade (CRISPR-associated complex for antiviral defense), a 405-kD complex consisting of 11 subunits of five Cas proteins ($Cse1_1$, $Cse2_2$, $Cas7_6$, $Cas5_1$, and $Cas6e_1$) and a 61-nucleotide (nt) crRNA. The crRNA consists of a 32-nt guide sequence flanked by 5′ and 3′ handles, the sequences of which are derived from the repeats (Fig. 1A). Cascade recognizes DNA as foreign if it contains a region complementary to the crRNA guide (protospacer) that is adjacent to a small 3–base pair (bp)

[1]Department of Biochemistry and Molecular Biology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21205, USA. [2]Photon Sciences Directorate, Brookhaven National Laboratory, Upton, NY 11973, USA.
***Corresponding author. E-mail: scott.bailey@jhu.edu**

sequence element called a protospacer adjacent motif (PAM). CRISPR arrays lack a PAM sequence, enabling discrimination between self and nonself DNA (*21*). Cascade binding to foreign DNA initiates at the PAM and then proceeds through base pairing between the crRNA and target strand, displacing the nontarget strand to produce an R loop (*14*). R-loop formation is unidirectional, proceeding 5′ to 3′ across the crRNA guide (*22*). Cryo–electron microscopy (cryo-EM) reconstructions of target-bound Cascade suggest that base pairing between the crRNA guide and the target strand occurs stepwise through a series of discontinuous helical segments to form the guide-target hybrid (*15*). The structure of these discontinuous segments is currently unknown, nor is it known how Cascade stabilizes this structure.

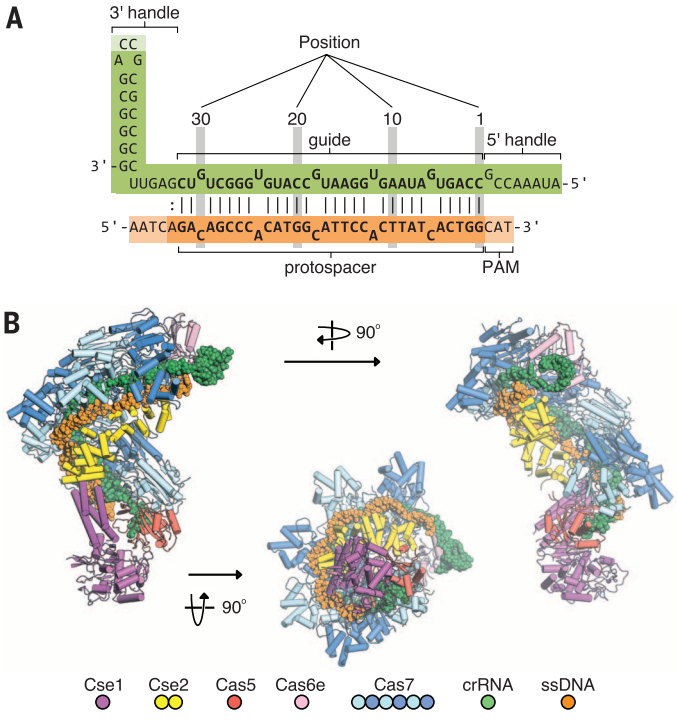

Fig. 1. Overall structure of ssDNA-bound Cascade. (**A**) Cartoon schematic of the crRNA and ssDNA target. Watson-Crick base pairing between the two strands is indicated by lines, two dots indicates an observed shear base pair. The position of each base pair within the target-guide hybrid is also indicated. (**B**) Three orthogonal views of ssDNA-bound Cascade showing the seahorse architecture of the complex. Protein subunits are shown as a ribbon representation in the colors indicated. The crRNA (green) and ssDNA target (orange) are displayed in a spheres representation.

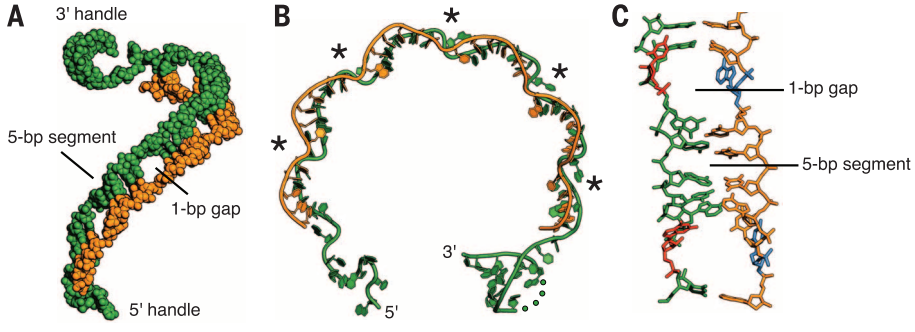

Fig. 2. Structure of the crRNA and ssDNA target. (**A**) Overall structure of the two nucleic acids, highlighting the ribbon-like structure of the guide-target hybrid. The crRNA (green) and ssDNA target (orange) are displayed in a spheres representation. (**B**) Ribbon representation of the crRNA and ssDNA revealing the periodic kinking (marked by asterisks) in the backbone of both strands. (**C**) Structure of a 5-bp segment and disrupted base pairs. Disrupted RNA and DNA nucleotides are colored red and blue, respectively. In all panels, the crRNA and ssDNA target are colored green and orange, respectively.

Crystal structure of the Cascade-ssDNA complex

To gain insights into the structural organization of Cascade and into target recognition, we have determined the crystal structure of a ~425-kD complex of *E. coli* Cascade bound to a complementary single-stranded DNA (ssDNA) target. We screened a variety of ssDNA and double-stranded DNA (dsDNA) targets for crystallization with Cascade. The best crystals contained a complex of Cascade and a 40-nt ssDNA target consisting of a 32-nt protospacer sequence, a 5′ PAM sequence (5′-CAT-3′), and a 3′ extension (5 nt) (Fig. 1A). The structure—containing the ssDNA target strand, crRNA, and all of the 11 Cascade subunits—was refined to an R_{work}/R_{free} ratio of 0.225/0.269 at 3.03 Å resolution (table S1), has good stereochemistry (table S1), and accounts for the vast majority of the expected residues (table S2). Although the ssDNA target contained a PAM and a 5-nt 3′ extension, only the protospacer and the adjacent 3′ nucleotide of the extension were visible in the electron density map.

The crystal structure of ssDNA-bound Cascade has the same seahorse architecture observed by cryo-EM (*14*, *15*) (Fig. 1B). The body is formed by a helical filament of six Cas7 subunits (Cas7.1 to 7.6) wrapped around the crRNA guide, with a head-to-tail dimer of Cse2 (Cse2.1 and Cse2.2) at the belly. Cas6e and the 3′ handle of crRNA cap the Cas7 filament at the head while Cas5 and the 5′ handle cap the tail. The N-terminal base of Cse1 is positioned at the tail of the filament; the C-terminal four-helix bundle contacts Cse2.2. The ssDNA target is juxtaposed to the guide region of the crRNA in a groove formed by the Cas7 filament, the four-helix bundle of Cse1, and the Cse2 dimer (Fig. 1B). The structures of the *E. coli* subunits are structurally similar to the stand-alone structures of their respective homologs (fig. S1) (*11*, *12*, *21*, *23*–*30*).

Two cryo-EM reconstructions of target-bound Cascade are available, one bound to a 32-nt ssRNA

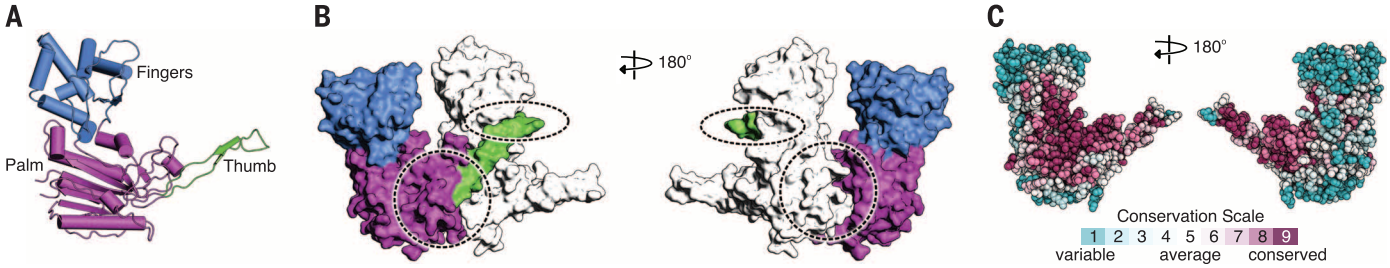

Fig. 3. Structure of the Cas7 subunit. (**A**) Ribbon representation of Cas7 colored by domain: thumb (green), fingers (blue), and palm (purple). (**B**) Surface representation of adjacent Cas7 subunits. The locations of the large and small interfaces are highlighted by dashed circles. (**C**) Surface representation of Cas7 colored according to evolutionary conservation.

target (*15*) and one bound to a 72-bp dsDNA target with a functional PAM sequence (*31*). The structural organization of Cascade in the two reconstructions is similar except for the position of the Cse1 base, which has been implicated in PAM binding (*21*). The crystal structure of Cascade fits best in the EM density of dsDNA-bound Cascade, requiring only small changes in the position of Cas6e and the 3′ handle for an optimal fit (fig. S2). Lattice contacts appear to hold the base of Cse1 in the dsDNA-bound conformation. Thus, the structure of ssDNA-bound Cascade observed in the crystal is comparable to that bound to dsDNA, the in vivo target of Cascade (*3*, *31*).

The guide-target hybrid adopts a ribbon conformation

The two strands of the guide-target hybrid do not twist around one another in a helix, but instead adopt an underwound ribbon-like structure reminiscent of a ladder (Fig. 2A). The 5′ and 3′ ends of the curved target strand are ~102 Å apart, roughly the length of straight B-form dsDNA with an identical sequence (~107 Å). Underwinding is facilitated by kinks that occur every sixth base pair in the backbone of both strands of the hybrid (Fig. 2B). At each kink, complementary nucleotides are rotated ~90°, in opposing directions, from the axis of the duplex (Fig. 2C). Thus, consistent with cryo-EM studies, the guide-target ribbon contains five 5-bp segments separated by 1-bp gaps (*15*). The remaining 2 nt of the guide region also form Watson-Crick base pairs with the target strand and are followed by a shear base pair between the nucleotide 5′ of the protospacer (A5) and the first nucleotide (G41) in the 5′ handle of the crRNA (fig. S3A). The importance of this shear base pair is currently unclear, as there is no reported conservation at this position of target sequence.

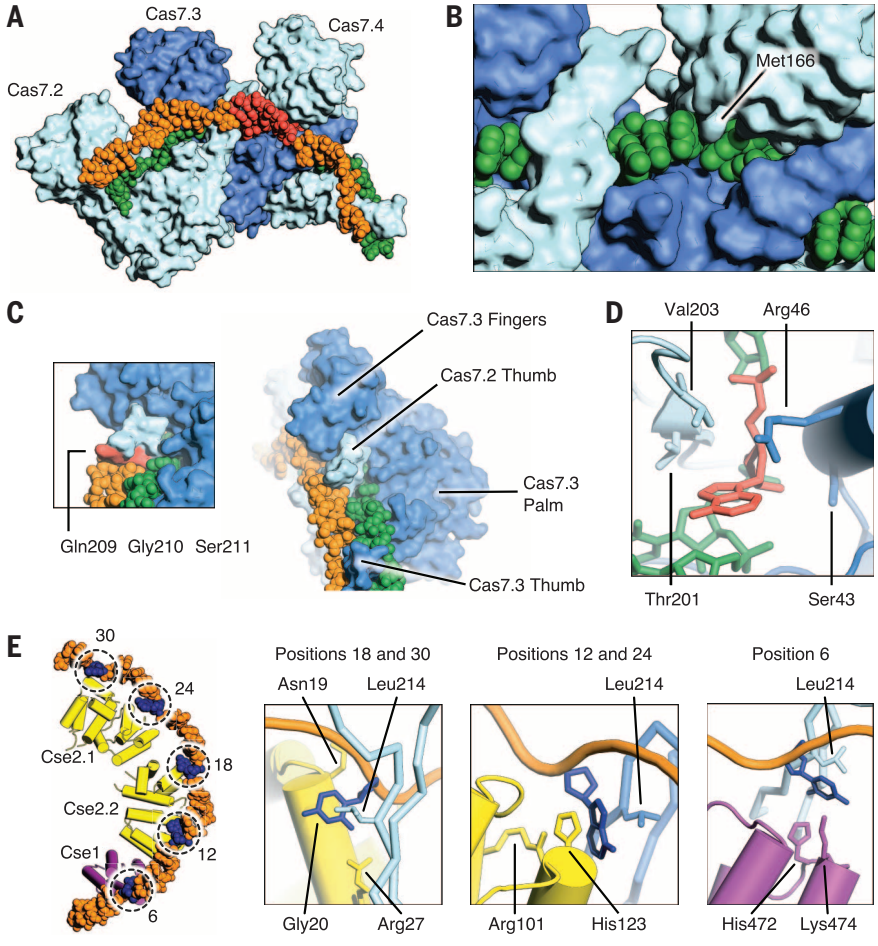

Fig. 4. Stabilization of the guide-target hybrid by Cas7, Cse1, and Cse2. (**A**) Surface representation showing the interaction of Cas7 subunits with the guide-target hybrid. The 5-bp segment bound between the fingers of Cas7.4 and the palm of Cas7.4 is colored red. (**B**) Close-up view of the bound crRNA. The DNA target has been removed for clarity. Intercalation by Met^{166} from Cas7 is highlighted. (**C**) Surface representation showing that the thumb of Cas7 pushes through the guide-target hybrid at the 1-bp gaps. Boxed is the interaction of thumb residues with the minor groove face of the hybrid. (**D**) Binding site for the disrupted RNA bases. (**E**) Left: Overview of the interactions between the ssDNA target and Cse2.1, Cse2.2, and Cse1. The proteins are represented as ribbons, the DNA as a surface. The positions of the disrupted DNA nucleotides (dark blue) are indicated. Right: Close-up views of the DNA nucleotide binding pockets. In all panels, the proteins and nucleic acids are colored as in Fig. 1.

The structure of each repeating unit, a 5-bp segment plus the two displaced nucleotides, is essentially identical, superposing to an average root mean square deviation (RMSD) of ~1.20 Å over 209 atoms (fig. S3B). The local conformation of the 5-bp segments is highly distorted from canonical A-form, owing to the nonuniform roll and twist angles between base pairs. However, this distortion is necessary for the hybrid to remain continuous throughout the guide region, as the geometry of an A-form duplex is incompatible with the spacing between kink sites (fig. S3C).

The observed structure of the guide-target hybrid predicts that mutation of the disrupted nucleotides would not affect binding by Cascade. Consistent with this, mutations at position 6 of the protospacer (Fig. 1A) have no effect on target binding, while mutations at positions 1 to 5 or 7 to 8 (Fig. 1A) greatly reduce affinity (*32*). In addition, recent high-throughput genetic experiments revealed that mutations at the sites of the disrupted bases are readily tolerated in vivo (*33*).

The Cas7 filament and its interaction with the head of Cascade

The structure of Cas7 resembles a right hand, consisting of fingers (residues 59 to 180), palm (residues 1 to 58, 181 to 189, and 224 to 263), and thumb (residues 190 to 223) domains (Fig. 3A). The palm contains the modified RNA-recognition motif (RRM) characteristic of many Cas protein families (*34*). The domain architecture and key surface features appear conserved among Cas7 family proteins (fig. S4), which suggests that the structure of the Cas7 filament and consequently the ribbon-like structure of the guide-target hybrid (see below) are likely conserved in other type I and type III CRISPR-Cas systems. Within Cascade, the six Cas7 subunits form a right-handed filament, with a pitch of ~135 Å, around the guide-target hybrid (Fig. 1B). The filament is arranged such that the thumb of one Cas7 subunit, composed of an extended β hairpin, extends toward the fingers of the adjacent subunit (Fig. 3B). The conformations of Cas7.2 to Cas7.6 are essentially identical, except that the fingers of Cas7.6 are disordered in the structure (fig. S5A). Three points of contact stabilize the Cas7 filament, one between Cas7 and the guide region of the crRNA (see below) and two involving conserved protein-protein contacts (Fig. 3, B and C) between Cas7 subunits. The larger of these contacts buries ~1500 $Å^2$ of solvent-accessible surface area per subunit and is formed by packing the base of the thumb and the back of the palm of one subunit against the front of the palm of the neighboring subunit (Fig. 3B). The second, smaller interface buries ~400 $Å^2$ of solvent-accessible surface area per subunit and is formed between the tip of the thumb of one subunit and the fingers of the neighboring subunit (Fig. 3B).

At the head of Cascade, the Cas7 filament and the 3′ end of the crRNA are capped by Cas6e (fig. S5B). This region appears to be flexible relative to the rest of Cascade, as the corresponding electron density is quite weak. One face of Cas6e engages the 3′ hairpin of the crRNA, as observed

with *Thermus thermophilus* Cas6e (*28*, *30*), while the opposite face binds the thumb of Cas7.1 (fig. S5C). Although a portion of the Cas7.1 thumb is disordered in the structure, it is clear that it has a conformation distinct to the other Cas7 subunits. Specifically, the Cas7 thumb rotates ~90° toward Cas6e, adopting an open loop structure (fig. S5D) that binds primarily with a conserved face of the C-terminal RRM of Cas6e (fig. S5C).

Subunit interactions with the guide-target hybrid

A highly interdigitated network of protein–nucleic acid interactions stabilizes the ribbon conformation of the guide-target hybrid (Fig. 4A). Extensive contacts between the guide region of the crRNA and the Cas7 filament bury a large portion of the crRNA backbone, leaving the bases solvent-exposed (Fig. 4B). The absence of direct contacts between protein side chains and bases of the crRNA explains the lack of sequence specificity by Cascade for the guide sequence. The target strand is bound primarily through Watson-Crick hydrogen bonding with the guide but also interacts with the Cas7, Cse1, and Cse2 subunits (Fig. 1B). Few interactions are seen with the DNA backbone, rendering the majority of the backbone solvent-exposed (Fig. 1B).

Each 5-bp segment of the hybrid is situated between the palm of one Cas7 subunit and the fingers of the adjacent subunit (Fig. 4A). Several highly conserved polar and positively charged residues (Arg^{20}, Lys^{27}, Ser^{40}, Gln^{42}, Lys^{45}, and Lys^{49}) from the palm of one Cas7 contact the RNA backbone, while the fingers from the adjacent Cas7 subunit (residues 109 to 111 and 163 to 169) contact both strands of the hybrid across the minor groove (Fig. 4C). In particular, the side chain of a highly conserved methionine (Met^{166}) partially intercalates with the bases at the third and fourth position of each RNA segment (Fig. 4B). As a result, the bases at these positions are separated, which helps distort each segment away from A-form geometry, facilitating the underwound structure of the crRNA guide.

The thumb of Cas7 both prohibits base pairing at the kink sites of the hybrid and stabilizes the conformation of the 5-bp segments. As the thumb of one Cas7 subunit extends toward the fingers of the adjacent subunit it passes directly between the strands of the guide-target hybrid, displacing the nucleotides and disrupting their base pairing (Fig. 4C). As a result, adjacent Cas7 subunits completely encircle the guide region of the crRNA (Fig. 4, A to C). Residues from the middle of the Cas7 thumb (Trp^{199}, Phe^{200}, Thr^{201}, Ala^{202}, Leu^{214}, and His^{213}) (fig. S6) also stabilize the position of the nucleotides that flank the ends of the 5-bp segments through van der Waals contacts. Residues at the tip of the thumb (Gln^{209}, Gly^{210}, and Ser^{211}) contact the bases of the target strand on the minor groove face of the duplex (Fig. 4C).

Distinct binding pockets accommodate the displaced RNA and DNA nucleotides, respectively. Each binding pocket is solvent-exposed and provides minimal contacts with the Watson-Crick edges of the bases, consistent with the lack of sequence specificity at these positions. Each displaced RNA nucleotide adopts the syn conformation, is similarly positioned above the backbone of the downstream RNA, and is contacted by residues from both the Cas7 palm (Ser^{43} and Arg^{46}) and thumb (Thr^{201} and Val^{203}) (Fig. 4D). Each displaced DNA nucleotide adopts the anti conformation, is sandwiched on one side by Leu^{214} from the nearest Cas7 thumb, and, depending on its position along the hybrid, by residues from Cse1, Cse2.1, or Cse2.2 (Fig. 4E). Cse1 contacts the first displaced DNA nucleotide (position 6) (Fig. 1A) with residues Lys^{474} and His^{472} from the four-helix bundle (Fig. 4E). This interaction is consistent with cryo-EM studies showing ordering of this region of Cse1 (residues 406 to 414) on target binding (*15*). The second and fourth displaced DNA nucleotides (positions 12 and 24) stack with His^{123} and contact Arg^{101} from either Cse2.2 or Cse2.1, respectively (Fig. 4E). The third and fifth displaced nucleotides (positions 18 and 30) make contact with Asn^{19}, Gly^{20}, and Arg^{27} from either Cse2.2 or Cse2.1, respectively (Fig. 4E). We also observe a salt bridge between Arg^{26} from Cse2.1 and Cse2.2 and backbone phosphates at positions 17 and 29. Consistent with these observations, the highly conserved Arg^{26}, Arg^{27}, and Arg^{101} from Cse2 have been implicated as important for DNA binding by electrophoretic mobility shift assays (*25*).

Interactions capping the tail of Cascade

Cas5 caps the tail of the Cas7 filament at the 5′ end of the crRNA (Fig. 5A). The structure of Cascade reveals that Cas5 is structurally related to Cas7, as it consists of a palm (residues 1 to 78 and 115 to 224) and a thumb (residues 79 to 114) domain, but lacks a fingers domain (Fig. 5B). The Cas5 palm contains an RRM that superposes with the palm of Cas7 (RMSD of ~3.3 Å between 123 Cα atoms). The thumb domains of Cas5 and Cas7 adopt a similar orientation with respect to their palm domains and appear to have similar functions within Cascade. The thumb of Cas5 displaces nucleotide G8 in the 5′ handle of the crRNA in a manner reminiscent of how the Cas7 thumb displaces RNA nucleotides in the guide. The thumb of Cas5 also provides similar residues (Leu^{89} and Thr^{87}) to the binding pocket of the G8 nucleotide, as the Cas7 thumb does for the displaced RNA nucleotides (Val^{203} and Thr^{201}). Additionally, two residues from the thumb of Cas5 (Tyr^{85} and Gln^{105}) stabilize the first duplex segment (positions 1 to 5) through van der Waals contacts with the exposed face of the base at position 1 of the hybrid (fig. S6).

The first seven nucleotides of the 5′ handle form a hook-like structure (Fig. 5C) that sits between Cas5 and Cas7.6 (Fig. 5D), forming extensive contacts between the RNA backbone and highly conserved charged and polar residues from both Cas5 (Arg^{25}, Ser^{34}, Arg^{48}, Arg^{49}, Arg^{148}, and Arg^{149}) and Cas7.6 (Lys^{45}, Arg^{49}, and Arg^{46}) (Fig. 5E). The first three nucleotides (A1, U2, and A3) are splayed and their bases sandwiched in three distinct pockets formed by conserved residues from Cas5 (fig. S7). Sequence-specific contacts are observed between nucleotide U2 and the backbone of Tyr^{145} (fig. S7). The bases of nucleotides A4, A5, and C6 form a triplet stack that sits orthogonal to nucleotide A3 (fig. S7). The backbone of this stack packs against the palm domains of both Cas5 and Cas7.6, leaving the bases solvent-exposed. The side chain of Arg^{206}

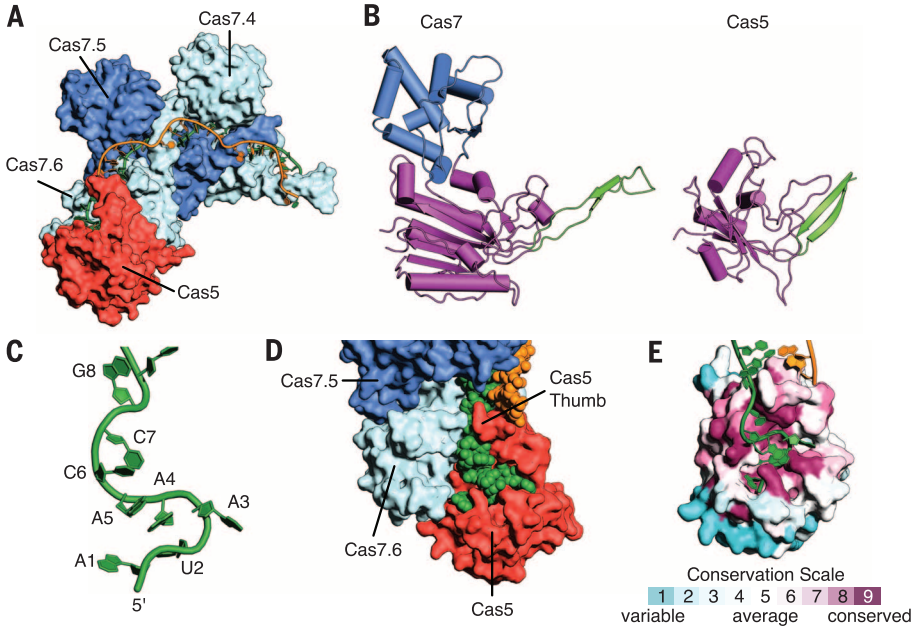

Fig. 5. Structure of Cas5 and the 5′ handle of the crRNA. (**A**) Surface representation of Cas5 capping the end of the Cas7 filament. (**B**) Ribbon representation of Cas7 and Cas5 colored by domain: thumb (green), fingers (blue), and palm (purple). (**C**) Structure of the 5′ hook of the crRNA. (**D**) Close-up view of the interaction between Cas5, Cas7.6, and the 5′ hook. (**E**) Surface representation of Cas5 colored according to evolutionary conservation, highlighting the conservation of the residues that bind the 5′ hook.

from the palm of Cas5 also stacks with the base of nucleotide A4 (fig. S7). Finally, the position of C7 is stabilized through sequence-specific hydrogen bonds with Arg^{108} from the Cas5 thumb (fig. S7). Thus, only nucleotides U2 and C7 form sequence-specific interactions with Cas5 in ssDNA-bound Cascade, consistent with sequence alignments showing these positions are the invariant in the 5′ handle (*35*).

Modeling interactions of Cascade, dsDNA target, and Cas3

Cryo-EM studies (*31*) revealed that Cascade binds the PAM-proximal end of a dsDNA target between Cas7.5, Cas7.6, and the L1 loop of Cse1, which has been linked to PAM binding (*21*). The fingers of Cas7.6 and the L1 loop of Cse1 are disordered in the crystal structure of ssDNA-bound Cascade, which suggests that these regions are mobile in the absence of dsDNA. Using the cryo-EM density as a guide, we modeled the position of this dsDNA and the fingers of Cas7.6 onto the crystal structure. The resulting model suggests interactions between the DNA backbone and a cluster of basic residues from the fingers of both Cas7.5 (Lys^{137}, Lys^{138}, and Lys^{141}) and Cas7.6 (His^{67} and Lys^{105}) (Fig. 6A). The orientation of the modeled duplex and the path of the target strand agree with atomic force microscopy studies suggesting that Cascade bends dsDNA upon binding (*18*).

R-loop formation by Cascade displaces the nontarget strand of the protospacer. Footprinting experiments have shown that the 3′ end but not the 5′ end of the displaced strand is protected by Cascade binding (*14*, *21*). The structure of ssDNA-bound Cascade reveals a prominent basic groove that could serve as a binding site for the 3′ end of the displaced strand (Fig. 6B). Spanning from Cas6e to the four-helix bundle of Cse1, the groove is lined with conserved basic residues from Cas7 (Lys^{34}, Lys^{299}, and Lys^{301}) and Cse2 (Arg^{53}, Arg^{142}, Arg^{143}, and Arg^{110}) (fig. S8A). Because the PAM is bound at the L1 loop of Cse1 (*31*), the displaced strand must first loop around the four-helix bundle of Cse1 to gain access to the groove, in agreement with footprinting data showing that the 5′ end of the displaced strand is exposed. After R-loop formation, Cascade recruits the Cas3 helicase-nuclease (*18*–*20*). Negative stain reconstruction of a complex between dsDNA-bound Cascade and Cas3 revealed Cas3 binds Cascade between the four-helix bundle and the base of Cse1 (*31*). Mapping this interaction onto the crystal structure reveals the residues that likely mediate Cas3 binding (fig. S8B); these include Glu^{192}, Glu^{280}, Asn^{376}, Thr^{383}, and two conserved loops located at the base of Cse1 (residues 288 to 294 and residues 318 to 323).

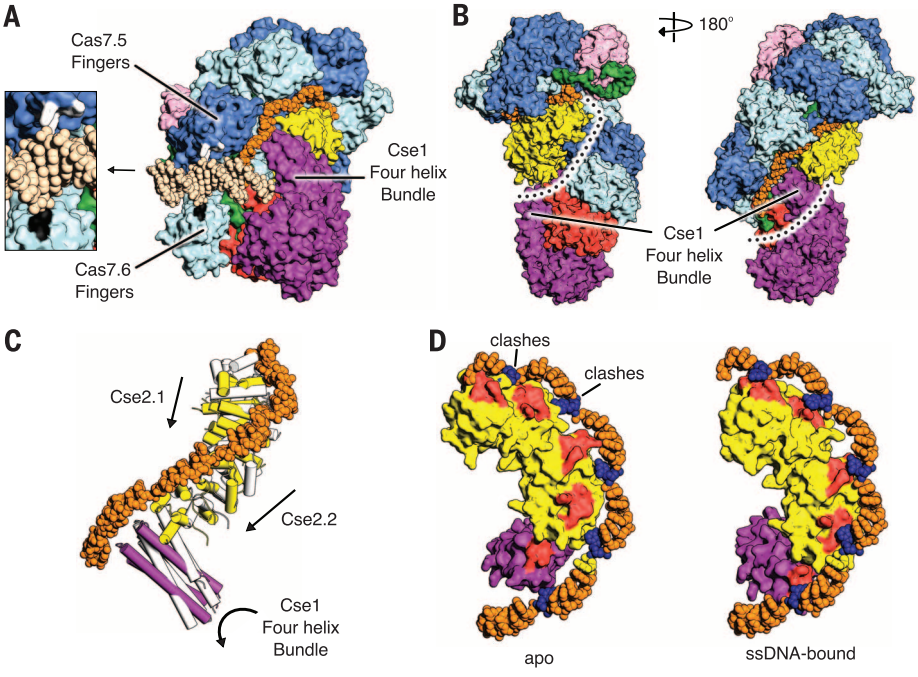

Fig. 6. Modeling interactions with dsDNA and associated conformation changes. (**A**) Surface representation of Cascade and the modeled dsDNA and Cas7.6 fingers domain. Boxed is a close-up view of the residues predicted to bind the dsDNA: Lys^{137}, Lys^{138}, and Lys^{141} from the fingers of Cas7.5 (white) and His^{67} and Lys^{105} from the fingers of Cas7.6 (black). (**B**) Surface representation of Cascade, with the proposed path of the nontarget strand indicated by a dotted line. (**C**) Ribbon representation of Cse2.1, Cse2.2, and the four-helix bundle of Cse1, highlighting the structural rearrangements that accompany ssDNA binding. The target strand is shown in orange. The subunits in the apo conformation are colored white; those in the ssDNA-bound conformation are colored as in Fig. 1. (**D**) Surface representations of Cse2.1, Cse2.2, and the four-helix bundle of Cse1 in ssDNA-bound (right) and apo (left) conformations. The binding pockets for displaced DNA nucleotides (dark blue) are colored red. Clashes between Cse2.1 and the distal end of the ssDNA target are indicated.

Cryo-EM studies have shown that target binding triggers structural rearrangements of the Cse1 and Cse2 subunits (*15*). To further understand these changes, we fit each subunit of the crystal structure of ssDNA-bound Cascade into the ~9 Å cryo-EM reconstruction of apo Cascade as a rigid body, except for Cse1, which required individual placement of its two domains (fig. S9). As previously established, superposition of the apo Cascade model with the structure of ssDNA-bound Cascade reveals a concerted conformational rearrangement. The two Cse2 subunits each move ~16 Å toward the four-helix bundle of Cse1, which subsequently undergoes a ~30° rotation away from Cse2.2 (Fig. 6C). We also observe a ~15° rotation of the base of Cse1 (*31*). Together, these rearrangements facilitate the formation of binding pockets for the displaced DNA nucleotides (Fig. 6D). Additionally, in the apo conformation, Cse2.1 blocks access to the distal end of the guide (positions 24 to 30) (Fig. 6D). Therefore, repositioning Cse2.1 facilitates base pairing in this region, consistent with previous binding data showing that short oligonucleotides complementary to the 3′ region of the crRNA guide have lower affinity than short oligonucleotides that bind to the 5′ region (*15*).

Implications for target binding and recruitment of Cas3

Analysis of the crystal structure, in the context of the reported literature, provides a model for how Cascade recognizes foreign DNA and recruits Cas3. During target recognition, Cascade searches dsDNA for potential PAM sites, possibly facilitated by sequence-independent interactions between the DNA backbone and basic residues from the fingers of Cas7.5 and Cas7.6 (Fig. 6A) (*31*). Once a PAM is found, the adjacent duplex is destabilized and base pairing between the crRNA guide and the target strand proceeds along the guide in 5-bp increments separated by 1-bp gaps (*15*, *22*, *32*).

Cascade incrementally binding 5-nt segments of the target strand likely increases the fidelity of target recognition, by a mechanism similar to that used by RecA (*36*–*38*). RecA catalyzes strand exchange in homologous recombination. Thus, despite differing cellular roles, Cascade and RecA both function by correctly pairing a single strand of RNA or DNA, respectively, with a complementary strand in a DNA duplex, respectively. Comparison of the structures of the guide-target hybrid in ssDNA-bound Cascade and the heteroduplex bound to the RecA filament (*36*) reveals globally similar distortions (fig. S10), as previously predicted (*2*). RecA is thought to bind its target in 3-nt increments to enforce high fidelity during strand exchange; if so, this suggests that Cascade may use a similar conformational proofreading mechanism (*36*–*38*). Specifically, conformational proofreading by Cascade involves optimizing the cost-benefit ratio of binding targets by introducing a structural difference between the crRNA guide and the foreign DNA. The energetics of target binding are the net sum of the energy gained by base pairing and the energy lost in conformational distortion (i.e., formation of the guide-target hybrid). Nonspecific targets lacking

favorable base pair interactions will not overcome the energetic costs associated with conformational distortion, whereas specific targets can do so. Thereby, enhanced specificity is achieved. Genetic and biochemical assays identified a seed region (positions 1 to 5 and 7 to 8) in crRNA required for high-affinity binding to a target (*32*). Consistent with this proposed conformational fidelity mechanism, mismatches in the seed region would inhibit formation of the initial 5-nt segments, thereby terminating binding.

Cascade undergoes structural rearrangement of its Cse1 and Cse2 subunits upon target binding. Propagation of base pairing between the crRNA guide and target strand across the target facilitates these conformation changes, which form the binding pockets for the disrupted DNA nucleotides (Fig. 6D). Moreover, movement of Cse2.1 relieves a steric block at the distal end of the guide (positions 24 to 30) (Fig. 6D), enabling base pairing in this region. Recent single-molecule studies monitoring DNA supercoiling revealed that Cascade binding to target DNA is unstable until base pairs form at the distal end of the guide (*22*). Our structural analysis suggests that base pairing in this region would prevent Cse2.1 from reassuming its apo position, effectively locking Cascade on the DNA. This locking mechanism could also act as an additional proofreading step, as targets that cannot form base pairs across the entire length of the guide will not be stably bound (*22*).

After target recognition, Cascade recruits the Cas3 helicase-nuclease, likely through interactions with two conserved loops at the base of Cse1 (fig. S8B) (*31*). Once recruited, Cas3 nicks the displaced nontarget strand ~7 to 11 nt from the 3′ end of the PAM (*19*, *20*), consistent with the predicted path of the nontarget strand (Fig. 6B and fig. S8A). Cas3 then loads onto the newly formed ssDNA end (*31*) and continues to progressively degrade the foreign DNA.

REFERENCES AND NOTES

1. R. Barrangou *et al.*, *Science* **315**, 1709–1712 (2007).
2. R. Sorek, C. M. Lawrence, B. Wiedenheft, *Annu. Rev. Biochem.* **82**, 237–266 (2013).
3. S. J. J. Brouns *et al.*, *Science* **321**, 960–964 (2008).
4. J. Carte, R. Wang, H. Li, R. M. Terns, M. P. Terns, *Genes Dev.* **22**, 3489–3496 (2008).
5. E. Deltcheva *et al.*, *Nature* **471**, 602–607 (2011).
6. K. S. Makarova *et al.*, *Nat. Rev. Microbiol.* **9**, 467–477 (2011).
7. C. R. Hale *et al.*, *Cell* **139**, 945–956 (2009).
8. C. Rouillon *et al.*, *Mol. Cell* **52**, 124–134 (2013).
9. J. van der Oost, E. R. Westra, R. N. Jackson, B. Wiedenheft, *Nat. Rev. Microbiol.* **12**, 479–492 (2014).
10. J. Zhang *et al.*, *Mol. Cell* **45**, 303–313 (2012).
11. K. H. Nam *et al.*, *Structure* **20**, 1574–1584 (2012).
12. N. G. Lintner *et al.*, *J. Biol. Chem.* **286**, 21643–21656 (2011).
13. N. Heidrich, J. Vogel, *Mol. Cell* **52**, 4–7 (2013).
14. M. M. Jore *et al.*, *Nat. Struct. Mol. Biol.* **18**, 529–536 (2011).
15. B. Wiedenheft *et al.*, *Nature* **477**, 486–489 (2011).
16. R. H. J. Staals *et al.*, *Mol. Cell* **52**, 135–145 (2013).
17. M. Spilman *et al.*, *Mol. Cell* **52**, 146–152 (2013).
18. E. R. Westra *et al.*, *Mol. Cell* **46**, 595–605 (2012).
19. T. Sinkunas *et al.*, *EMBO J.* **32**, 385–394 (2013).
20. S. Mulepati, S. Bailey, *J. Biol. Chem.* **288**, 22184–22192 (2013).
21. D. G. Sashital, B. Wiedenheft, J. A. Doudna, *Mol. Cell* **46**, 606–615 (2012).
22. M. D. Szczelkun *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 9798–9803 (2014).
23. S. Mulepati, A. Orr, S. Bailey, *J. Biol. Chem.* **287**, 22445–22449 (2012).
24. Y. Agari, S. Yokoyama, S. Kuramitsu, A. Shinkai, *Proteins* **73**, 1063–1067 (2008).
25. K. H. Nam, Q. Huang, A. Ke, *FEBS Lett.* **586**, 3956–3961 (2012).
26. Y. Koo, D. Ka, E.-J. Kim, N. Suh, E. Bae, *J. Mol. Biol.* **425**, 3799–3810 (2013).
27. A. Hrle *et al.*, *RNA Biol.* **10**, 1670–1678 (2013).
28. D. G. Sashital, M. Jinek, J. A. Doudna, *Nat. Struct. Mol. Biol.* **18**, 680–687 (2011).
29. A. Ebihara *et al.*, *Protein Sci.* **15**, 1494–1499 (2006).
30. E. M. Gesner, M. J. Schellenberg, E. L. Garside, M. M. George, A. M. Macmillan, *Nat. Struct. Mol. Biol.* **18**, 688–692 (2011).
31. M. L. Hochstrasser *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 6618–6623 (2014).
32. E. Semenova *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 10098–10103 (2011).
33. P. C. Fineran *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **111**, E1629–E1638 (2014).
34. K. S. Makarova, N. V. Grishin, S. A. Shabalina, Y. I. Wolf, E. V. Koonin, *Biol. Direct* **1**, 7 (2006).
35. V. Kunin, R. Sorek, P. Hugenholtz, *Genome Biol.* **8**, R61 (2007).
36. Z. Chen, H. Yang, N. P. Pavletich, *Nature* **453**, 489–4 (2008).
37. S. C. Kowalczykowski, *Nature* **453**, 463–466 (2008).
38. Y. Savir, T. Tlusty, *Mol. Cell* **40**, 388–396 (2010).

ACKNOWLEDGMENTS

We thank R. McMacken, B. Learn, J. Berger, D. Leahy, and J. Kavran for helpful discussions; J. Bosch for providing the tungsten clusters used in the soaking experiments; I. Mathews for help with data collection; and J. Kavran for critical reading of the manuscript. Supported by NIH grant GM097330 (S.B). Data for this study were measured at beamline X25 of the National Synchrotron Light Source (NSLS) and at beamlines 7-1, 11-1, and 12-2 of the Stanford Synchrotron Radiation Lightsource (SSRL). Funding for X25 comes principally from the Offices of Biological and Environmental Research and of Basic Energy Sciences of the U.S. Department of Energy (DOE) and from National Center for Research Resources grant P41RR012408 and NIH grant P41GM103473. Use of the SSRL is supported by the DOE Office of Science, Office of Basic Energy Sciences, under contract DE-AC02-76SF00515. The SSRL Structural Molecular Biology Program is supported by the DOE Office of Biological and Environmental Research and by NIH grant P41GM103393. The atomic coordinates and structure factors have been deposited into the Protein Data Bank with the accession code 4QYZ.

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/345/6203/1479/suppl/DC1
Materials and Methods
Figs. S1 to S10
Tables S1 and S2
References (*39*–*47*)

5 June 2014; accepted 4 August 2014
Published online 14 August 2014;
10.1126/science.1256996

REPORTS

QUANTUM GASES

Observation of Fermi surface deformation in a dipolar quantum gas

K. Aikawa,[1] S. Baier,[1] A. Frisch,[1] M. Mark,[1] C. Ravensbergen,[1,2] F. Ferlaino[1,2]*

In the presence of isotropic interactions, the Fermi surface of an ultracold Fermi gas is spherical. Introducing anisotropic interactions can deform the Fermi surface, but the effect is subtle and challenging to observe experimentally. Here, we report on the observation of a Fermi surface deformation in a degenerate dipolar Fermi gas of erbium atoms. The deformation is caused by the interplay between strong magnetic dipole-dipole interaction and the Pauli exclusion principle. We demonstrate the many-body nature of the effect and its tunability with the Fermi energy. Our observation provides a basis for future studies on anisotropic many-body phenomena in normal and superfluid phases.

The Fermi-liquid theory, formulated by Landau in the late 1950s, is one of the most powerful tools in modern condensed-matter physics (*1*). It captures the behavior of interacting Fermi systems in the normal phase, such as electrons in metals and liquid ^{3}He (*2*). Within this theory, the interaction is accounted by dressing the fermions as quasi-particles with an effective mass and an effective interaction. The ground state is the so-called Fermi sea, in which the quasi-particles fill one-by-one all the states up to the Fermi momentum, k_F. The Fermi surface (FS), which separates occupied from empty states in k-space, is a sphere of radius k_F for isotropically interacting fermions in uniform space. The FS is crucial for understanding system excitations and Cooper pairing in superconductors. When complex interactions act, the FS can get modified. For instance, strongly correlated electron systems violate the Fermi-liquid picture, giving rise to a deformed FS, which spontaneously breaks the rotational invariance of the system (*3*). Symmetry-breaking FSs have been studied in connection with electronic liquid crystal phases (*4*) and Pomeranchuk instability (*5*) in solid-state systems. Particularly relevant is the nematic phase, in which anisotropic behaviors spontaneously emerge and the system acquires an orientational order, while preserving its translational invariance (*3*).

[1]Institut für Experimentalphysik and Zentrum für Quantenphysik, Universität Innsbruck, Technikerstraße 25, 6020 Innsbruck, Austria. [2]Institut für Quantenoptik und Quanteninformation, Österreichische Akademie der Wissenschaften, 6020 Innsbruck, Austria.
***Corresponding author. E-mail: francesca.ferlaino@uibk.ac.at**