Adaptation processes that build CRISPR immunity – creative destruction, updated.

Chun H. Lau, Ryan Reeves and Edward L. Bolt*

School of Life Sciences, University of Nottingham. U.K.
Correspondence to ed.bolt@nottingham.ac.uk

Abstract
Prokaryotes can defend themselves against invading mobile genetic elements (MGEs) by acquiring immune memory against them. The memory is a DNA database located at specific chromosomal sites called CRISPRs (clustered regularly interspaced short palindromic repeats) that store fragments of MGE DNA. These are utilised to target and destroy returning MGEs, preventing re-infection. The effectiveness of CRISPR-based immune defence depends on "adaptation" reactions that *capture* and *integrate* MGE DNA fragments into CRISPRs. This provides the means for immunity to be delivered against MGEs in "interference" reactions. Adaptation and interference are catalysed by Cas (CRISPR-associated) proteins, aided by enzymes well known for other roles in cells. We survey the molecular biology of CRISPR adaptation, highlighting entirely new developments that may help us to understand how MGE DNA is captured. We focus on processes in *Escherichia coli*, punctuated with reference to other prokaryotes that illustrate how common requirements for adaptation, DNA capture and integration, can be achieved in different ways. We also comment on how CRISPR adaptation enzymes, and their antecedents, can be utilised for biotechnology.

Summary points

- Prokaryotic CRISPR-Cas systems record encounters with mobile genetic elements (MGEs) by storing MGE DNA fragments in CRISPR loci. This provides them with immune memory that can be utilised to destroy returning MGEs.
- "Adaptation" creates CRISPR DNA arrays from nuclease destruction of MGEs.
- Cas1 and Cas2 proteins catalyse adaptation, aided by other proteins; we summarise new evidence that RecBCD nuclease activity is not a factor in adaptation, but that it's helicase activity is, and suggest new possible roles for DnaK and Cas1-Cas2.
- Cas1 proteins are diverse in form and function but retain common core enzyme activities. These can be utilised for synthetic biology.

CRISPR DNA – the keystone of prokaryotic adaptive immunity

Horizontal gene transfer (HGT) − the spread of genes into a cell from an alien cell − is frequent in prokaryotes, driven by mobile genetic elements (MGEs) that are commonly plasmids, transposons and viruses. MGEs are abundant and have a huge impact on ecology, evolution and animal health

- this is evident from the spread of antimicrobial resistance in many contexts, for example in healthcare (1). But to avoid metabolic burden and cell death that may also be associated with MGEs, prokaryotes have also evolved barriers to them that take multiple forms and which include CRISPR-Cas systems (*herein* CRISPR systems). These give prokaryotes adaptive immunity against an MGE that is re-encountered after immunity has been established against it. Similarly, CRISPR systems impede gene flow that arises from inter-species mating in halophilic archaea (2).

CRISPR systems are highly diverse, exemplified by regularly updated molecular phylogenies; a recent analysis is presented in (3). Common principles underpin all CRISPR systems, most obviously the centring of immunity on specialized chromosomal sites called CRISPRs (clustered regularly interspaced short palindromic repeats), see Figure 1. CRISPRs are a repository for DNA fragments that have been captured and integrated by Cas1-Cas2 enzymes, detailed more below. The appearance of CRISPRs was first described in (4), as an array of 29 base-pair DNA repeats separated by variable 32 base-pair "spacings", observations that coalesced into the naming of CRISPR and their CRISPR-associated (Cas) genes (5) – the biographical details are covered in a recent review (6).

The sources of the variable DNA "spacer" sequences within CRISPR is the key to immunity – spacers from an MGE provide immune protection, but spacers are also derived from the host cell chromosome. Indeed, the factors determining selection of DNA targets that ultimately become a new spacer is an active area of research into CRISPR systems. Data mining of spacer sequences across CRISPR systems identified the source for 7% of spacers, averaged across all systems, of which 80-90% were from viruses, although the 93% spacer "dark matter" showed sequence characteristics suggesting MGE origin (7). When spacers are transcribed from CRISPR DNA to RNA (crRNA) they target Cas proteins to MGEs, in processes called "Interference", which include nucleolytic destruction of the MGE: More details about interference reactions and how they have been utilised for genetic editing, are covered elsewhere (8-12). Therefore CRISPRs containing tens of different spacers can be utilised by a cell to target multiple sequences in multiple MGEs, providing immune protection by degrading MGE nucleic acid. By expansion of CRISPR loci through acquisition of new spacers the CRISPR system ensures stable inheritance of acquired adaptive genetic change, resembling Lamarckian evolution.

CRISPR adaptation – processes that build the immune memory

The effectiveness of CRISPR systems depends on adaptation, processes that create CRISPR arrays by adding new spacer sequences. However, spacers are also lost from CRISPR arrays as part of the evolutionary pressures exerted on prokaryotes by new or prevailing MGEs, reviewed in (13). This CRISPR dynamics chimes with processes described in economic theory as "creative destruction"; "*industrial mutation that incessantly revolutionizes…incessantly destroying the old one, incessantly creating a new one"*(14). Figure 1 gives an overview of how adaptation expands CRISPR arrays by adding new DNA sequences.

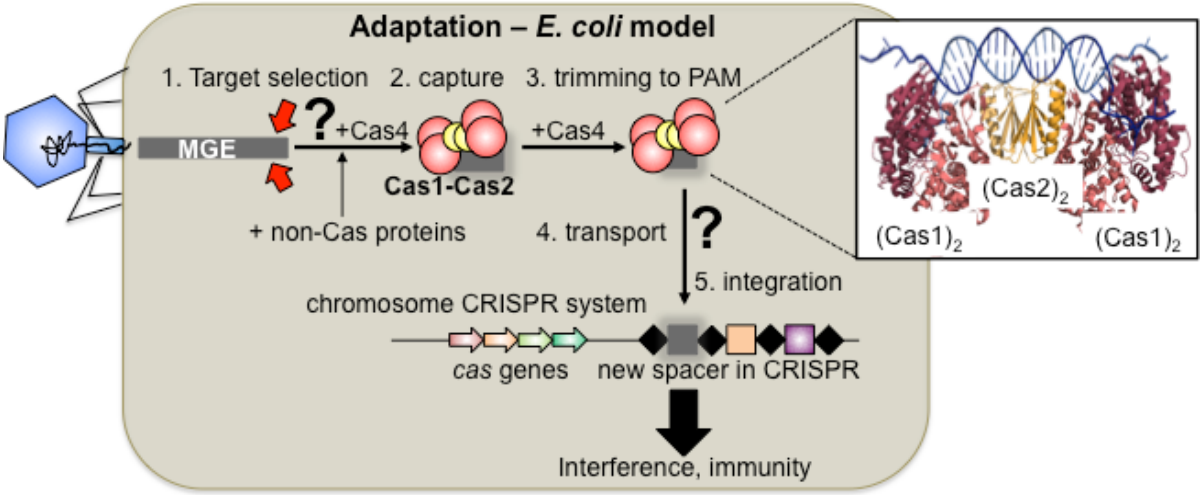

Figure 1 – *Summary of how CRISPR-Cas immunity is built by adaptation, focussing on the E. coli model system, but additionally highlighting how aspects of DNA target selection and capture are not understood (denoted by?), and the importance of Cas4 protein for DNA capture and protospacer processing at PAM sequences in other CRISPR systems – see main text for details. The Cas1-Cas2 complex catalyses integration of a new spacer into CRISPR by a well defined mechanism that is dependent on DNA duplex positioned within a dimer of Cas1 dimers (red) that sandwiches a Cas2 dimer (yellow)(15,16). DNA capture and Integration are aided by additional nucleic acid processing enzymes of the host cell that are not part of CRISPR-Cas systems – see (17-21) and main text for details.*

Adaptation involves multiple nucleic acid processing events; DNA target selection, DNA capture with further processing at the freed DNA ends, and DNA integration into CRISPR. Adaptation against an unfamiliar MGE, to which the cell has no prior immunity, generates a spacer from it for the first time. This is called *naïve* adaptation – it establishes immunity against an MGE *de novo*. Subsequent nuclease activities during interference reactions against the returning MGE generate more MGE DNA or RNA fragments for capture. This bolsters immunity and updates it; an important aspect of this is to counteract MGE mutations that escape from interference, resulting in lost immunity: This form of adaptation is called "priming"(22).

Cas1 and Cas2 proteins are essential for DNA capture and integration into CRISPR systems (23), forming a complex that has been detailed at atomic resolution (15,16,24,25), and has revealed mechanism for integration into CRISPR (26,27). In the Cas1-Cas2 adaptation complex of *E. coli* two Cas1 dimers sandwich a Cas2 dimer, altogether holding a 23bp DNA duplex with five-nucleotide ssDNA tails. This becomes a 33-bp protospacer following nucleophilic integration into CRISPR DNA and DNA gap repair generates a new DNA repeat (28). It is not known how Cas1-Cas2 loaded with captured protospacer DNA ready for integration arrives at CRISPR loci but once it is there other factors help to ensure that integration is achieved so that a spacer can be utilised for interference (20,21,29-31).

DNA target selection and capture establishes CRISPR immunity *de novo*.

DNA capture by Cas1-Cas2 complex is a pre-requisite for generating and updating immunity against an MGE, but its molecular details are one of the least understood aspects of CRISPR immunity. DNA capture by Cas1-Cas2 during primed adaptation at interference reactions, in *E. coli* cells that have already established immunity to an MGE, relies on the Cas3 nuclease-translocase (32-34) and the captured DNA is single-stranded (35). This intriguing finding suggests that a mechanism is needed for conversion of captured ssDNA to dsDNA prior to integration. Other questions remain about DNA capture; for example, in naïve adaptation, when interference reactions are absent because there is no pre-existing immunity, how does Cas1-Cas2 identify "non-self" MGE DNA over "self" host DNA? In this context, how and in what form is DNA excised and captured as a fragment for integration? How do functional interactions between non-Cas proteins and Cas1-Cas2 help to achieve adaptation, and how are CRISPR systems functionally integrated more generally with cellular physiology? This last question is highlighted by involvement of the Cas3 interference protein in biofilm formation in *Pseudomonas* bacteria (36,37), and potential links between CRISPR systems and many other non-Cas genes (38). A recent review article covered adaptation in detail (39); below we highlight developments in understanding DNA capture, including significant new information:

(a) *RecBCD helicase, not nuclease, is needed for adaptation* − The RecBCD protein complex targets DNA ends, which trigger its nuclease-helicase activities either upon binding dsDNA breaks in the chromosome or MGE, or upon binding the ends of linear MGEs. The bimodal nuclease activities of RecBCD have been prominent in models of DNA capture during naïve adaption (17,18) − details of RecBCD activities in DNA degradation, repair and recombination can be found in reference (40). In adaptation, DNA fragments generated by RecBCD nuclease activities are proposed to be captured by Cas1-Cas2. Furthermore, nuclease activities of RecBCD are substantially moderated on encountering chromosomal *Chi* DNA sequences, thereby reducing the availability of DNA fragments for capture by Cas1-Cas2. This effect of *Chi* is crucial for the naïve adaptation model because it offers a mechanism for adaptation to target "non-self" MGE DNA over "self" chromosomal DNA (17); MGEs largely lack *Chi* sequences and so would be more

voraciously degraded by RecBCD acting as a nuclease. This would therefore present MGE DNA as capture substrates for Cas1-Cas2 more frequently than host cell DNA. However, recent analysis of RecBCD in *E. coli* indicates that *RecBCD nuclease activities are not important for naïve adaptation*, revealed by genetic dissection of the several different activities of RecBCD in cells (19). This study also identified that residual helicase activity of RecBC, without the RecD 5' - 3' helicase subunit, is important for supporting adaptation. Inactivation of the nuclease activity of RecB in RecBC, by utilizing the allele *recB1080*, also had little effect on adaptation further supporting that RecBCD nuclease activity overall is not necessary for adaptation (19). Additionally, a 2011 study using *E. coli* Cas1 protein had noted its interaction with RecB and RecC proteins (41). Using microscopy we can visualise Cas1 foci in *E. coli* cells, which increase in intensity when RecBC, but not RecD are present in cells (Christian Rudolph, Brunel University, pers. comm.). These developments suggest that current models of how RecBCD assists adaptation in bacterial cells may need updating. Archaea lack RecBCD, or its functional homologue AddAB, but achieve the same resection of DNA ends to promote recombination through activities of analogous enzymes − it will be interesting to see in archaea if those are also important for CRISPR adaptation, perhaps via a common mechanism requiring availability of DNA ends.

(b) *Cas1-Cas2 as a nuclease in DNA capture* − In addition to integrase activity targeted to CRISPR loci, Cas1 proteins are nucleases on a wide variety of DNA substrates *in vitro* (41-43), consistent with integrase enzymes more generally (44). Cas1-Cas2 integration reactions that deliver a new spacer into CRISPR depend on availability of 3' ssDNA ends in the captured protospacer DNA, for nucleophilic attack on CRISPR DNA (28,29,45). In *E. coli* and *P. atrosepticum*, nuclease activity of Cas1 exposes 3' ssDNA 'tails' by chewing back the DNA ends after capture (16,46). The nuclease activity may also facilitate DNA target selection and capture in adaptation, potentially providing an all-in-one activity that may not require other cellular nucleases, at least for naïve adaptation (19,46). Nuclease activities of Cas1 reported in recent work (19) required formation of the Cas1-Cas2 complex, and were not supported by Cas1 in isolation. It also indicated that Cas1-Cas2 is proficient as a nuclease at DNA ends, in line with free DNA ends stimulating DNA capture for adaptation (17,47,48).

For CRISPR immunity to be effective newly integrated spacers must be functional for interference reactions that deliver the immune response. This boils down to delivery of a spacer sequence to invading MGEs as CRISPR RNA (crRNA), within a ribonucleoprotein complex. In this way base pairing between crRNA spacer and cognate MGE DNA or RNA leads to nucleolytic destruction of the MGE (33,49,50). Proteins within interference complexes detect the MGE through a protospacer adjacent motif (PAM) DNA sequence, establishing effective interference. Interference in some CRISPR systems, including *E. coli*, utilises the Cas8 subunit of Cascade ribonucleoprotein complex to detect a PAM 5'-AAG-3' (51), triggering R-loop formation and recruitment of Cas3 nuclease to the MGE DNA (50,52-54). This is reflected in integration reactions, which help to

ensure functionality of new spacers for interference reactions by intrinsic sequence preferences for Cas1 at the spacer-repeat boundary in CRISPR (29,31,45,55-57).

However, the sequence preferences of Cas1-Cas2 at integration cannot ensure specificity for MGE DNA during target selection and capture by Cas1-Cas2, because PAM sequences will be more frequent within the host cell genome than MGEs. The ability to distinguish "non-self" DNA for capture instead of "self" DNA is elegantly achieved in primed adaptation, in which cells already have immunity, by recruitment of Cas1-Cas2 to Cascade and Cas3 interference reactions at MGEs (52). In naïve adaptation this is not possible, and other factors may be needed to achieve a distinction between non-self and self nucleic acids: Involvement of RecBCD and *Chi* in this has already been highlighted above. It is also intriguing that Cas1 is part of the DnaK interactome (58), and physical interaction between DnaK and Cas1-Cas2 is observed in our laboratory. The significance of this to adaptation, if any, arises from the requirement for DnaK as a molecular chaperone in the early stages of DNA replication by bacteriophage and conjugative plasmids, which could position Cas1-Cas2 for MGE target selection during naïve adaptation. Alternatively, perhaps there is no need for self/non-self discrimination by Cas1-Cas2 during naïve adaptation; instead in response to concerted attack from MGEs there is a "spacer storm" in which Cas1-Cas2 in each afflicted cell captures as much DNA as rapidly as possible from any source, MGE or chromosomal, achieving immunity against the MGE in some cells and auto-immunity and thus cell death in other cells, that ensures survival of the population as a whole.

(d) *Cas4, a RecB-like nuclease, supports adaptation* − Cas4, a sequence homologue of the RecB nuclease, is distributed in CRISPR systems across bacteria and archaea. Cas4 nuclease activities ensure correct processing of captured DNA at PAM sequences, and orientate protospacers at integration in ways that ensure spacers are functional for interference (21,59-62). Cas4 achieves this with Cas1 and Cas2 proteins in DNA processing machines that ready DNA for integration at CRISPRs.

(e) *Adaptation by Cas1-Cas2 complexes that capture RNA* − diversity in CRISPR systems is further evidenced by their association with reverse transcriptase enzymes in several phylogenetic clades (63). In some of these (e.g. *Marinomonas mediterranea*) Cas1 is fused to a reverse transcriptase (RT) allowing capture of RNA protospacers that become DNA spacers (64). *In vitro* characterisation of this RT-Cas1 showed that the protein in complex with Cas2 ligated both DNA and RNA protospacers into the CRISPR array by the first nucleophilic attack on CRISPR. The *M. mediterranea* RT-Cas1-Cas2 complex catalysed cDNA synthesis from the ligated RNA template generating a DNA spacer. In addition, this RT-Cas1 protein contains an N-terminal Cas6 domain that processes pre-crRNA in readiness for assembly into interference reactions (65). The discovery of these RT-Cas1 proteins is also driving new developments in biotechnology, described more below.

The roots of adaptation − Casposase, Cas1 and simplified DNA integration

Cas1 protein is a universal marker that a CRISPR system is functional at creating immunity by capture and integration of new spacer DNA. However, two groups of Cas1 homologues are not associated with CRISPR systems (66), one within a superfamily of DNA transposons, called 'casposons' (67). Purified casposon Cas1 homologue is an integrase, resulting in its naming 'Casposase' (68). CRISPR systems and Cas1 protein are thought to have originated from casposons (69), and Casposase catalyses integration of casposon DNA in a similar way to spacer integration by the Cas1-Cas2 complex in CRISPR systems, but with reduced complexity (Figure 2). Casposase from the archaeon *Acidulprofundum boonei* creates a target site duplication (TSD) upon casposon integration that resembles the CRISPR repeat duplication during adaptation. It was shown that Casposase biased the integration into a target site consisting of a TSD and at least 18 bp upstream of the TSD. The upstream 18-bp segment encodes the TΨC loop of a tRNA gene, which determines the precise location of integration, therefore acting analogously to the leader sequence upstream of CRISPR arrays (70).

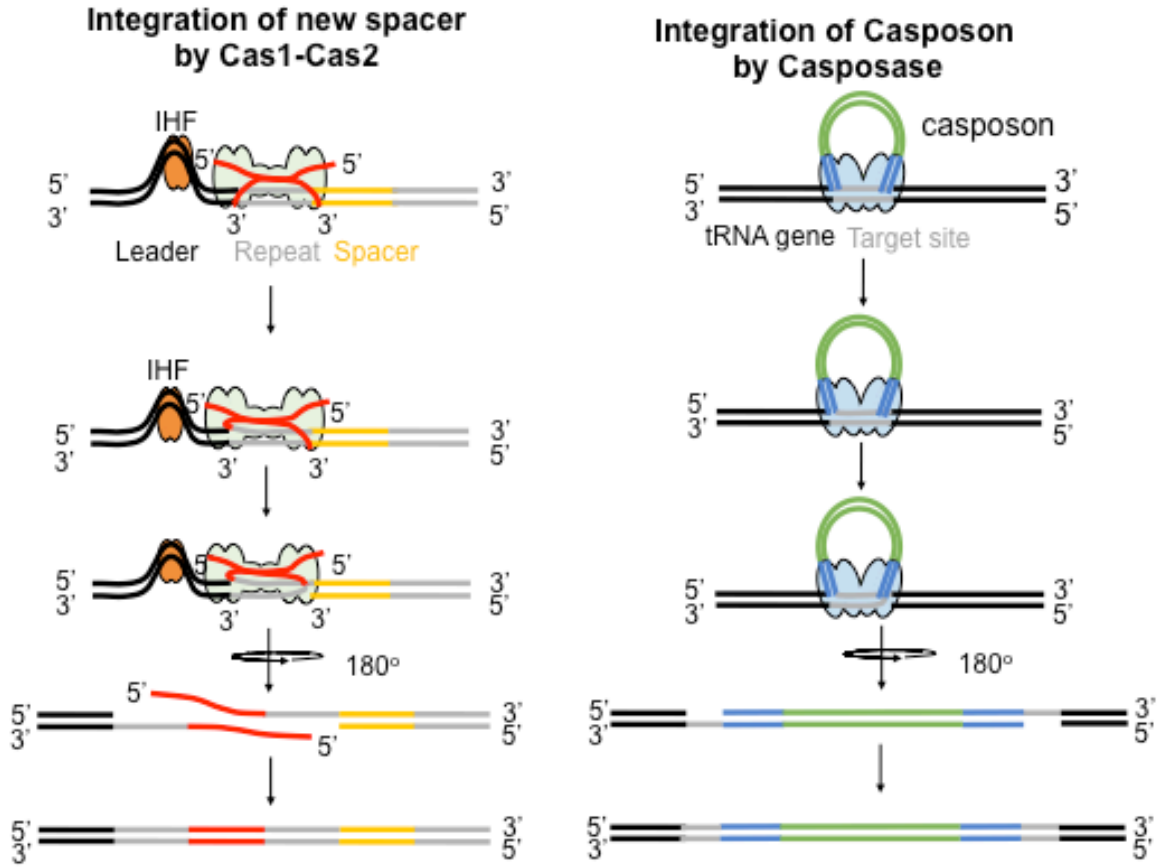

Figure 2 - *Comparison of DNA spacer integration by Cas1-Cas2 complex (left), and Casposase-catalysed integration of a casposon (right). Spacer integration by Cas1-Cas2 complex is directed to the leader-proximal repeat helped by integration host factor (IHF) in some bacterial CRISPR. For casposon integration, Casposase is recruited at a target site downstream of a tRNA gene.*

Cas1-Cas2 catalyses the first nucleophilic attack by the 3'-OH group of the spacer at the CRISPR leader-repeat boundary, and Casposase by the 3'-OH group of the casposon at the tRNA-target site boundary. The second nucleophilic attack occurs at the opposite strand of the target DNA distal to the leader or tRNA gene. Resulting single strand DNA gaps are thought to be repaired by gap filling polymerase (18) resulting in CRISPR repeat or target site duplication.

Casposase targets DNA that is flanked by terminal inverted repeats (TIRs), giving specificity for capture and integration of casposons into target DNA. The molecular basis for this specificity at TIRs is not clear, but might reside in a conserved C-terminal helix-turn-helix (HTH) domain that is absent from Cas1 proteins. Resemblance between Casposase and Cas1 encoded by CRISPR systems highlights the evolution of CRISPR systems from casposons. A likely intermediate form between Casposase and Cas1-Cas2 was recently identified (71). This Cas1 protein is from a CRISPR system that lacks Cas2 protein, and forms tetramers to catalyse new spacer integration. These Cas1 and Casposase proteins, like RT-Cas1, have potential in genetic editing biotechnologies.

Adaptation in biotechnology – genetic editing by Cas1 - Cas2 and related proteins

Proteins catalysing CRISPR Interference have been adapted in multiple ways for genetic engineering. Most prevalent is Cas9 from *Streptococcus* species, which can be modified and fused to other proteins to achieve gene disruptions and replacements, gene regulation, single base editing and genome visualization (72). The Cas1-Cas2 adaptation complex has also been utilised for biotechnologies, especially in using its DNA capture and integration activities for making DNA-based molecular recordings. In natural CRISPR systems the Cas1-Cas2 catalysed integration of new spacers occurs at the leader-proximal end of the CRISPR array, upstream of older spacers. This generates a chronological record of encounters with mobile genetic elements (MGEs). The *E. coli* Cas1-Cas2 complex is most efficient at integrating synthetic oligonucleotides into the CRISPR array *in vitro* and *in vivo*. By harnessing these features, researchers demonstrated molecular recordings by supplying a series of artificial spacers over time (73). By DNA data storage in the form of Cas1-Cas2 integrated oligonucleotides it was possible to encode digital images and a short movie into the CRISPR array of a population of cells (74). These 'proofs-of-principle' using Cas1-Cas2 in DNA editing relies on precise integration activity at CRISPRs, and has potential for extended use when more is known about how Cas1-Cas2 and it's homologues capture DNA, or RNA. Molecular recording of RNA in *E. coli* cells has been successful by expressing RT-Cas1-Cas2 from *Fusicatenibacter saccharivorans*, an anaerobic bacterium present in the human gut (75). This "Record-seq" method could record, as DNA in CRISPR loci, the relative abundance of RNA in cells using the known transcriptomic data of oxidative and acid stress responses in *E. coli* as a benchmark. This could provide a new and powerful tool to understand more about regulation of gene expression under any physiological conditions.

Precise genetic “knock-in” of a user-defined DNA molecule without substantial DNA sequence homology is still a challenging task in genetic editing: Targeted integration of user-defined DNA through CRISPR-Cas9/Cas12a editing requires host cell homology-directed repair (HDR) systems, which are complex, of low efficacy in this context, and provoke additional genome instability, reviewed in (76). Various strategies have been developed to increase HDR efficiency, for example, by inhibiting alternative non-homologous end joining (NHEJ) pathways thereby stimulating HDR, and by fusing CtIP, a HDR protein, to Cas9 (77). Targeted integration can also be achieved by site-specific recombination catalysed by recombinase or transposase enzymes, but may be limited by strict target site requirements of the transposase. A transposase fused to a zinc-finger DNA binding protein was effective at integrating DNA into a site 6-17 bp from the zinc-finger domain binding site, but integration also required a TA dinucleotide target (78). As noted above, the integration activity of Cas1-Cas2, and natural fusion of Cas1 to reverse transcriptase have both been utilised for genetic editing, suggesting that Cas1 use in this regard could be extended into new editing tools. A recently described Cas1 protein can integrate 16-17 bp DNA fragments in the absence of Cas2 protein, with weak specificity for targeting the CRISPR array (71). Fusing this Cas1 homologue to Cas9 or deactivated Cas9 (dCas9) has potential as a RNA-guided DNA integrase. This same idea also has potential for development using Casposase instead of Cas1. Casposase can integrate DNA of up to 2.8 kb into another DNA molecule, at a site that is not constrained by any Casposase target sequence (68). Therefore creating Cas9-RNA-guided Cas1 DNA integrases may allow DNA editing by insertion of user-defined DNA sequences without requirement for HDR.

Concluding comments

Cas1-Cas2 build CRISPR arrays from DNA that is broken or destroyed, but we have incomplete understanding of the molecular details that trigger initial DNA target selection leading to its capture as a protospacer. The *E. coli* CRISPR system remains a useful model for studying those events by genetic and biochemical analysis of interactions between CRISPR systems and other non-CRISPR host cell systems, such as DNA repair and recombination. There is potential for Cas1 and its ancestral forms to be utilised more extensively in biotechnology, at least to assess whether they can be used to overcome technical challenges of genome editing reliant on homology-directed DNA repair.

References

1. Partridge, S.R., Kwong, S.M., Firth, N. and Jensen, S.O. (2018) Mobile Genetic Elements Associated with Antimicrobial Resistance. *Clin Microbiol Rev*, **31**.
2. Turgeman-Grott, I., Joseph, S., Marton, S., Eizenshtein, K., Naor, A., Soucy, S.M., Stachler, A.E., Shalev, Y., Zarkor, M., Reshef, L. *et al.* (2019) Pervasive acquisition of CRISPR memory driven by inter-species mating of archaea can limit gene transfer and influence speciation. *Nature microbiology*, **4**, 177-186.
3. Makarova, K.S., Wolf, Y.I. and Koonin, E.V. (2018) Classification and Nomenclature of CRISPR-Cas Systems: Where from Here? *CRISPR J*, **1**, 325-336.

4. Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. and Nakata, A. (1987) Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in Escherichia coli, and identification of the gene product. *J Bacteriol*, **169**, 5429-5433.
5. Jansen, R., Embden, J.D., Gaastra, W. and Schouls, L.M. (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol*, **43**, 1565-1575.
6. Mojica, F.J.M. and Montoliu, L. (2016) On the Origin of CRISPR-Cas Technology: From Prokaryotes to Mammals. *Trends Microbiol*, **24**, 811-820.
7. Shmakov, S.A., Sitnik, V., Makarova, K.S., Wolf, Y.I., Severinov, K.V. and Koonin, E.V. (2017) The CRISPR Spacer Space Is Dominated by Sequences from Species-Specific Mobilomes. *mBio*, **8**.
8. Wang, H.L.R., M.; Qi, L.S. (2016) CRISPR/Cas in genome editing and beyond. *Annual Reviews of Biochemistry*, **85**, 227 - 264.
9. Marraffini, L.A. and Sontheimer, E.J. (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science*, **322**, 1843-1845.
10. Marraffini, L.A. and Sontheimer, E.J. (2010) CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nature reviews. Genetics*, **11**, 181-190.
11. Knott, G.J. and Doudna, J.A. (2018) CRISPR-Cas guides the future of genetic engineering. *Science*, **361**, 866-869.
12. Hille, F., Richter, H., Wong, S.P., Bratovic, M., Ressel, S. and Charpentier, E. (2018) The Biology of CRISPR-Cas: Backward and Forward. *Cell*, **172**, 1239-1259.
13. Westra, E.R., van Houte, S., Gandon, S. and Whitaker, R. (2019) The ecology and evolution of microbial CRISPR-Cas adaptive immune systems. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, **374**, 20190101.
14. Schumpeter, J.A. (1994) *Capitalism, Socialism, Democracy.* Routledge (Verlag).
15. Nunez, J.K., Harrington, L.B., Kranzusch, P.J., Engelman, A.N. and Doudna, J.A. (2015) Foreign DNA capture during CRISPR-Cas adaptive immunity. *Nature*.
16. Wang, J., Li, J., Zhao, H., Sheng, G., Wang, M., Yin, M. and Wang, Y. (2015) Structural and Mechanistic Basis of PAM-Dependent Spacer Acquisition in CRISPR-Cas Systems. *Cell*.
17. Levy, A., Goren, M.G., Yosef, I., Auster, O., Manor, M., Amitai, G., Edgar, R., Qimron, U. and Sorek, R. (2015) CRISPR adaptation biases explain preference for acquisition of foreign DNA. *Nature*, **520**, 505-510.
18. Ivancic-Bace, I., Cass, S.D., Wearne, S.J. and Bolt, E.L. (2015) Different genome stability proteins underpin primed and naive adaptation in E. coli CRISPR-Cas immunity. *Nucleic acids research*, **43**, 10821-10830.
19. Radovcic, M., Killelea, T., Savitskaya, E., Wettstein, L., Bolt, E.L. and Ivancic-Bace, I. (2018) CRISPR-Cas adaptation in Escherichia coli requires RecBCD helicase but not nuclease activity, is independent of homologous recombination, and is antagonized by 5' ssDNA exonucleases. *Nucleic acids research*.
20. Nunez, J.K., Bai, L., Harrington, L.B., Hinder, T.L. and Doudna, J.A. (2016) CRISPR Immunological Memory Requires a Host Factor for Specificity. *Molecular cell*, **62**, 824-833.
21. Rollie, C., Graham, S., Rouillon, C. and White, M.F. (2018) Prespacer processing and specific integration in a Type I-A CRISPR system. *Nucleic acids research*, **46**, 1007-1020.
22. Datsenko, K.A., Pougach, K., Tikhonov, A., Wanner, B.L., Severinov, K. and Semenova, E. (2012) Molecular memory of prior infections activates the CRISPR/Cas adaptive bacterial immunity system. *Nature communications*, **3**, 945.
23. Yosef, I., Goren, M.G. and Qimron, U. (2012) Proteins and DNA elements essential for the CRISPR adaptation process in Escherichia coli. *Nucleic acids research*, **40**, 5569-5576.
24. Richter, C., Gristwood, T., Clulow, J.S. and Fineran, P.C. (2012) In vivo protein interactions and complex formation in the Pectobacterium atrosepticum subtype I-F CRISPR/Cas System. *PloS one*, **7**, e49549.
25. Nunez, J.K., Kranzusch, P.J., Noeske, J., Wright, A.V., Davies, C.W. and Doudna, J.A. (2014) Cas1-Cas2 complex formation mediates spacer acquisition during CRISPR-Cas adaptive immunity. *Nat Struct Mol Biol*, **21**, 528-534.
26. Wright, A.V., Liu, J.J., Knott, G.J., Doxzen, K.W., Nogales, E. and Doudna, J.A. (2017) Structures of the CRISPR genome integration complex. *Science*, **357**, 1113-1118.

27. Xiao, Y., Ng, S., Nam, K.H. and Ke, A. (2017) How type II CRISPR-Cas establish immunity through Cas1-Cas2-mediated spacer integration. *Nature*, **550**, 137-141.
28. Arslan, Z., Hermanns, V., Wurm, R., Wagner, R. and Pul, U. (2014) Detection and characterization of spacer integration intermediates in type I-E CRISPR-Cas system. *Nucleic acids research*, **42**, 7884-7893.
29. Rollie, C., Schneider, S., Brinkmann, A.S., Bolt, E.L. and White, M.F. (2015) Intrinsic sequence specificity of the Cas1 integrase directs new spacer acquisition. *Elife*, **4**.
30. Yoganand, K.N., Sivathanu, R., Nimkar, S. and Anand, B. (2017) Asymmetric positioning of Cas1-2 complex and Integration Host Factor induced DNA bending guide the unidirectional homing of protospacer in CRISPR-Cas type I-E system. *Nucleic acids research*, **45**, 367-381.
31. Moch, C., Fromant, M., Blanquet, S. and Plateau, P. (2017) DNA binding specificities of Escherichia coli Cas1-Cas2 integrase drive its recruitment at the CRISPR locus. *Nucleic acids research*, **45**, 2714-2723.
32. Kunne, T., Kieper, S.N., Bannenberg, J.W., Vogel, A.I., Miellet, W.R., Klein, M., Depken, M., Suarez-Diez, M. and Brouns, S.J. (2016) Cas3-Derived Target DNA Degradation Fragments Fuel Primed CRISPR Adaptation. *Molecular cell*, **63**, 852-864.
33. Brouns, S.J., Jore, M.M., Lundgren, M., Westra, E.R., Slijkhuis, R.J., Snijders, A.P., Dickman, M.J., Makarova, K.S., Koonin, E.V. and van der Oost, J. (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science*, **321**, 960-964.
34. Musharova, O., Vyhovskyi, D., Medvedeva, S., Guzina, J., Zhitnyuk, Y., Djordjevic, M., Severinov, K. and Savitskaya, E. (2018) Avoidance of Trinucleotide Corresponding to Consensus Protospacer Adjacent Motif Controls the Efficiency of Prespacer Selection during Primed Adaptation. *mBio*, **9**.
35. Musharova, O., Klimuk, E., Datsenko, K.A., Metlitskaya, A., Logacheva, M., Semenova, E., Severinov, K. and Savitskaya, E. (2017) Spacer-length DNA intermediates are associated with Cas1 in cells undergoing primed CRISPR adaptation. *Nucleic acids research*, **45**, 3297-3307.
36. Cady, K.C. and O'Toole, G.A. (2011) Non-Identity-Mediated CRISPR-Bacteriophage Interaction Mediated via the Csy and Cas3 Proteins. *J Bacteriol*, **193**, 3433-3445.
37. Cady, K.C., Bondy-Denomy, J., Heussler, G.E., Davidson, A.R. and O'Toole, G.A. (2012) The CRISPR/Cas adaptive immune system of Pseudomonas aeruginosa mediates resistance to naturally occurring and engineered phages. *J Bacteriol*, **194**, 5728-5738.
38. Shmakov, S.A., Makarova, K.S., Wolf, Y.I., Severinov, K.V. and Koonin, E.V. (2018) Systematic prediction of genes functionally linked to CRISPR-Cas systems by gene neighborhood analysis. *Proceedings of the National Academy of Sciences of the United States of America*, **115**, E5307-E5316.
39. McGinn, J. and Marraffini, L.A. (2019) Molecular mechanisms of CRISPR-Cas spacer acquisition. *Nature reviews. Microbiology*, **17**, 7-12.
40. Dillingham, M.S. and Kowalczykowski, S.C. (2008) RecBCD enzyme and the repair of double-stranded DNA breaks. *Microbiol Mol Biol Rev*, **72**, 642-671, Table of Contents.
41. Babu, M., Beloglazova, N., Flick, R., Graham, C., Skarina, T., Nocek, B., Gagarinova, A., Pogoutse, O., Brown, G., Binkowski, A. *et al.* (2011) A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. *Mol Microbiol*, **79**, 484-502.
42. Wiedenheft, B., Zhou, K., Jinek, M., Coyle, S.M., Ma, W. and Doudna, J.A. (2009) Structural basis for DNase activity of a conserved protein implicated in CRISPR-mediated genome defense. *Structure*, **17**, 904-912.
43. He, Y., Wang, M., Liu, M., Huang, L., Liu, C., Zhang, X., Yi, H., Cheng, A., Zhu, D., Yang, Q. *et al.* (2018) Cas1 and Cas2 From the Type II-C CRISPR-Cas System of Riemerella anatipestifer Are Required for Spacer Acquisition. *Frontiers in cellular and infection microbiology*, **8**, 195.
44. Yang, W. (2011) Nucleases: diversity of structure, function and mechanism. *Quarterly Reviews of Biophysics*, **44**, 1-93.
45. Nunez, J.K., Lee, A.S., Engelman, A. and Doudna, J.A. (2015) Integrase-mediated spacer acquisition during CRISPR-Cas adaptive immunity. *Nature*, **519**, 193-198.
46. Fagerlund, R.D., Wilkinson, M.E., Klykov, O., Barendregt, A., Pearce, F.G., Kieper, S.N., Maxwell, H.W.R., Capolupo, A., Heck, A.J.R., Krause, K.L. *et al.* (2017) Spacer capture

and integration by a type I-F Cas1-Cas2-3 CRISPR adaptation complex. *Proceedings of the National Academy of Sciences of the United States of America*, **114**, E5122-E5128.
47. Shiimori, M., Garrett, S.C., Chambers, D.P., Glover, C.V.C., 3rd, Graveley, B.R. and Terns, M.P. (2017) Role of free DNA ends and protospacer adjacent motifs for CRISPR DNA uptake in Pyrococcus furiosus. *Nucleic acids research*, **45**, 11281-11294.
48. Modell, J.W., Jiang, W. and Marraffini, L.A. (2017) CRISPR-Cas systems exploit viral DNA injection to establish and maintain adaptive immunity. *Nature*, **544**, 101-104.
49. Sinkunas, T., Gasiunas, G., Fremaux, C., Barrangou, R., Horvath, P. and Siksnys, V. (2011) Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *Embo J*, **30**, 1335-1342.
50. Jore, M.M., Lundgren, M., van Duijn, E., Bultema, J.B., Westra, E.R., Waghmare, S.P., Wiedenheft, B., Pul, U., Wurm, R., Wagner, R. *et al.* (2012) Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat Struct Mol Biol*, **18**, 529-536.
51. Sashital, D.G., Wiedenheft, B. and Doudna, J.A. (2012) Mechanism of foreign DNA selection in a bacterial adaptive immune system. *Molecular cell*, **46**, 606-615.
52. Redding, S., Sternberg, S.H., Marshall, M., Gibb, B., Bhat, P., Guegler, C.K., Wiedenheft, B., Doudna, J.A. and Greene, E.C. (2015) Surveillance and Processing of Foreign DNA by the Escherichia coli CRISPR-Cas System. *Cell*, **163**, 854-865.
53. Szczelkun, M.D., Tikhomirova, M.S., Sinkunas, T., Gasiunas, G., Karvelis, T., Pschera, P., Siksnys, V. and Seidel, R. (2014) Direct observation of R-loop formation by single RNA-guided Cas9 and Cascade effector complexes. *Proceedings of the National Academy of Sciences of the United States of America*, **111**, 9798-9803.
54. Westra, E.R., van Erp, P.B., Kunne, T., Wong, S.P., Staals, R.H., Seegers, C.L., Bollen, S., Jore, M.M., Semenova, E., Severinov, K. *et al.* (2012) CRISPR Immunity Relies on the Consecutive Binding and Degradation of Negatively Supercoiled Invader DNA by Cascade and Cas3. *Molecular cell*, **46**, 595-605.
55. Deveau, H., Barrangou, R., Garneau, J.E., Labonte, J., Fremaux, C., Boyaval, P., Romero, D.A., Horvath, P. and Moineau, S. (2008) Phage response to CRISPR-encoded resistance in Streptococcus thermophilus. *J Bacteriol*, **190**, 1390-1400.
56. Swarts, D.C., Mosterd, C., van Passel, M.W. and Brouns, S.J. (2012) CRISPR Interference Directs Strand Specific Spacer Acquisition. *PloS one*, **7**, e35888.
57. Wei, Y., Chesne, M.T., Terns, R.M. and Terns, M.P. (2015) Sequences spanning the leader-repeat junction mediate CRISPR adaptation to phage in Streptococcus thermophilus. *Nucleic acids research*, **43**, 1749-1758.
58. Calloni, G., Chen, T., Schermann, S.M., Chang, H.C., Genevaux, P., Agostini, F., Tartaglia, G.G., Hayer-Hartl, M. and Hartl, F.U. (2012) DnaK functions as a central hub in the E. coli chaperone network. *Cell reports*, **1**, 251-264.
59. Zhang, J., Kasciukovic, T. and White, M.F. (2012) The CRISPR associated protein Cas4 Is a 5' to 3' DNA exonuclease with an iron-sulfur cluster. *PloS one*, **7**, e47232.
60. Lemak, S., Nocek, B., Beloglazova, N., Skarina, T., Flick, R., Brown, G., Joachimiak, A., Savchenko, A. and Yakunin, A.F. (2014) The CRISPR-associated Cas4 protein Pcal_0546 from Pyrobaculum calidifontis contains a [2Fe-2S] cluster: crystal structure and nuclease activity. *Nucleic acids research*, **42**, 11144-11155.
61. Lee, H., Zhou, Y., Taylor, D.W. and Sashital, D.G. (2018) Cas4-Dependent Prespacer Processing Ensures High-Fidelity Programming of CRISPR Arrays. *Molecular cell*, **70**, 48-59 e45.
62. Shiimori, M., Garrett, S.C., Graveley, B.R. and Terns, M.P. (2018) Cas4 Nucleases Define the PAM, Length, and Orientation of DNA Fragments Integrated at CRISPR Loci. *Molecular cell*, **70**, 814-824 e816.
63. Toro, N., Martinez-Abarca, F., Gonzalez-Delgado, A. and Mestre, M.R. (2018) On the Origin and Evolutionary Relationships of the Reverse Transcriptases Associated With Type III CRISPR-Cas Systems. *Frontiers in microbiology*, **9**, 1317.
64. Silas, S., Mohr, G., Sidote, D.J., Markham, L.M., Sanchez-Amat, A., Bhaya, D., Lambowitz, A.M. and Fire, A.Z. (2016) Direct CRISPR spacer acquisition from RNA by a natural reverse transcriptase-Cas1 fusion protein. *Science*, **351**, aad4234.

65. Mohr, S.E., Hu, Y., Ewen-Campen, B., Housden, B.E., Viswanatha, R. and Perrimon, N. (2016) CRISPR guide RNA design for research applications. *The FEBS journal*, **283**, 3232-3238.
66. Makarova, K.S., Wolf, Y.I. and Koonin, E.V. (2013) The basic building blocks and evolution of CRISPR-CAS systems. *Biochem Soc Trans*, **41**, 1392-1400.
67. Krupovic, M., Makarova, K.S., Forterre, P., Prangishvili, D. and Koonin, E.V. (2014) Casposons: a new superfamily of self-synthesizing DNA transposons at the origin of prokaryotic CRISPR-Cas immunity. *BMC Biol*, **12**, 36.
68. Hickman, A.B. and Dyda, F. (2015) The casposon-encoded Cas1 protein from Aciduliprofundum boonei is a DNA integrase that generates target site duplications. *Nucleic acids research*, **43**, 10576-10587.
69. Koonin, E.V. and Krupovic, M. (2015) Evolution of adaptive immunity from transposable elements combined with innate immune systems. *Nature reviews. Genetics*, **16**, 184-192.
70. Beguin, P., Charpin, N., Koonin, E.V., Forterre, P. and Krupovic, M. (2016) Casposon integration shows strong target site preference and recapitulates protospacer integration by CRISPR-Cas systems. *Nucleic acids research*, **44**, 10367-10376.
71. Wright, A.V., Wang, J.Y., Burstein, D., Harrington, L.B., Paez-Espino, D., Kyrpides, N.C., Iavarone, A.T., Banfield, J.F. and Doudna, J.A. (2019) A Functional Mini-Integrase in a Two-Protein-type V-C CRISPR System. *Molecular cell*, **73**, 727-737 e723.
72. Adli, M. (2018) The CRISPR tool kit for genome editing and beyond. *Nature communications*, **9**, 1911.
73. Shipman, S.L., Nivala, J., Macklis, J.D. and Church, G.M. (2016) Molecular recordings by directed CRISPR spacer acquisition. *Science*, **353**, aaf1175.
74. Shipman, S.L., Nivala, J., Macklis, J.D. and Church, G.M. (2017) CRISPR-Cas encoding of a digital movie into the genomes of a population of living bacteria. *Nature*, **547**, 345-349.
75. Schmidt, F., Cherepkova, M.Y. and Platt, R.J. (2018) Transcriptional recording by CRISPR spacer acquisition from RNA. *Nature*, **562**, 380-385.
76. Cubbon, A., Ivancic-Bace, I. and Bolt, E.L. (2018) CRISPR-Cas immunity, DNA repair and genome stability. *Bioscience reports*, **38**.
77. Charpentier, M., Khedher, A.H.Y., Menoret, S., Brion, A., Lamribet, K., Dardillac, E., Boix, C., Perrouault, L., Tesson, L., Geny, S. *et al.* (2018) CtIP fusion to Cas9 enhances transgene integration by homology-dependent repair. *Nature communications*, **9**, 1133.
78. Feng, X., Bednarz, A.L. and Colloms, S.D. (2010) Precise targeted integration by a chimaeric transposase zinc-finger fusion protein. *Nucleic acids research*, **38**, 1204-1216.