Review

CellPress

Harnessing CRISPR–Cas systems for bacterial genome editing

Kurt Selle[1,2] and Rodolphe Barrangou[1]

[1] Department of Food, Bioprocessing, and Nutrition Sciences, North Carolina State University, Raleigh, NC 27606, USA
[2] Genomic Sciences Graduate Program, North Carolina State University, Raleigh, NC 27695, USA

Manipulation of genomic sequences facilitates the identification and characterization of key genetic determinants in the investigation of biological processes. Genome editing via clustered regularly interspaced short palindromic repeats (CRISPR)–CRISPR-associated (Cas) constitutes a next-generation method for programmable and high-throughput functional genomics. CRISPR–Cas systems are readily reprogrammed to induce sequence-specific DNA breaks at target loci, resulting in fixed mutations via host-dependent DNA repair mechanisms. Although bacterial genome editing is a relatively unexplored and underrepresented application of CRISPR–Cas systems, recent studies provide valuable insights for the widespread future implementation of this technology. This review summarizes recent progress in bacterial genome editing and identifies fundamental genetic and phenotypic outcomes of CRISPR targeting in bacteria, in the context of tool development, genome homeostasis, and DNA repair.

CRISPR–Cas systems and bacterial genome editing

Bacteria harbor CRISPR (see Glossary) and *cas* genes, which constitute an RNA-guided adaptive immune system against invasive genetic elements [1]. CRISPR–Cas-mediated immunity hinges on the distinct molecular processes of acquisition, expression, and interference [2]. Acquisition occurs via molecular 'sampling' of foreign DNA, from which short sequences, termed spacers, are integrated in a polarized manner at the leader end of a CRISPR array [1]. CRISPR arrays are transcribed constitutively and inducibly as directed by promoter elements in the preceding leader sequence during expression [3–5]. The transcript is processed selectively at each repeat sequence, forming mature CRISPR RNAs (crRNAs) that serve as small interfering RNAs (siRNAs). crRNAs guide Cas proteins for the sequence-specific recognition and cleavage of target DNA complementary to the spacer to effect interference. CRISPR–Cas systems encode universal *cas1* and *cas2* genes and are categorized as Type I, Type II, or Type III based on signature genes contributing to the distinct mechanisms by which each system confers interference [6]. Type I systems achieve immunity via the CRISPR-associated complex for antiviral defense (Cascade) through single-strand DNA nickase and exonuclease activity and are defined by the presence of Cas3 [7]. Features unique to Type II systems include the signature double-stranded (ds) DNA endonuclease Cas9, ancillary *trans*-activating crRNA (tracrRNA), and the biogenesis of crRNAs by RNase III [8,9]. Type III systems are marked by the signature gene Cas10 but are mechanistically diverse and less well defined, with some systems even capable of targeting RNA instead of DNA [10]. Delineation of CRISPR–Cas systems into 11 subtypes is similarly based on the presence of specific accessory *cas* genes and their respective genetic organization [6].

Cas9 effects interference in Type II systems through sequence-directed endonucleolysis at the target locus, achieved by concerted RuvC and HNH nickase activity [11,12]. The streamlined and multifunctional nature of Cas9 is practical for programmable genome editing in diverse organisms, requiring expression of only its cognate tracrRNA and a crRNA corresponding to the target sequence (Figure 1). The tipping point for this methodology was the creation of a single guide RNA (sgRNA) chimera that combines the functions of the native crRNA and tracrRNA duplex [13]. Cas9-mediated genome editing is programmable through the design of sgRNAs. The specificity of chromosomal cleavage hinges on the selection of a spacer sequence unique to the target allele and is further

Glossary

Alternative end joining (A-EJ): a Ku-independent DSB repair mechanism functioning through the annealing and ligation of microhomologous sequences.
Clustered regularly interspaced short palindromic repeats (CRISPR): a DNA locus comprising nearly identical repeats flanking sequences typically of exogenous origin.
CRISPR-associated (Cas): proteins that perform the three stages of CRISPR–Cas activity.
Double-stranded DNA break (DSB): DNA damage in which the phosphate backbone of both strands is hydrolyzed or cleaved.
Nonhomologous end joining (NHEJ): a Ku/ligase-dependent, error-prone DSB repair mechanism in bacteria and eukaryotes.
Protospacer: a target sequence in either the host genome or a foreign genetic element, such as a plasmid or phage, that is identical to a CRISPR spacer sequence.
Protospacer adjacent motif (PAM): a short (2–5 nt) conserved sequence that must be proximate to the target protospacer and is essential for acquisition and interference.
Seed: 8–12 bp of spacer proximate to the PAM, directly involved in crRNA::target DNA hybridization.
Single guide RNA (sgRNA): a single synthetic chimera combining the functions of native crRNA and tracrRNA.
Spacer: short sequences intervening between repeat sequences that correspond to target genomes.
***Trans*-activating CRISPR RNA (tracrRNA):** Type II-specific ancillary RNA that hybridizes to crRNAs and drives Cas9 activity.

Corresponding author: Barrangou, R. (rbarran@ncsu.edu).
Keywords: genome editing; CRISPR–Cas; Cas9.

0966-842X/
 http://dx.doi.org/10.1016/j.tim.2015.01.008

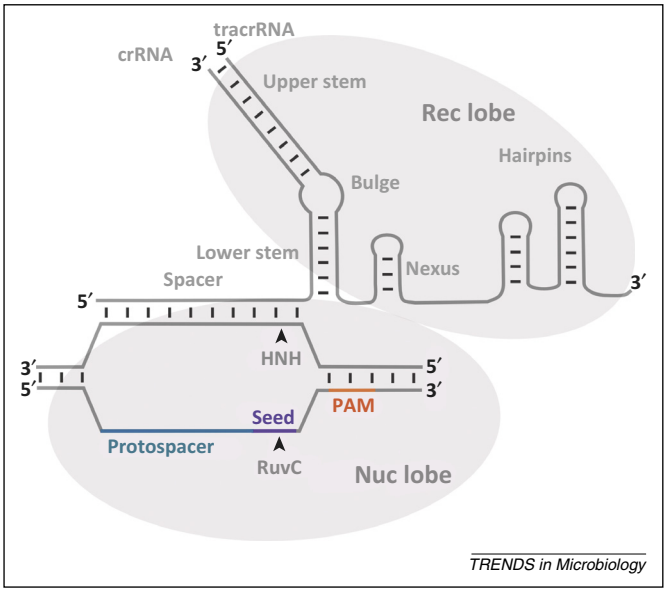

Figure 1. Clustered regularly interspaced short palindromic repeats (CRISPR)–CRISPR-associated 9 (Cas9) targeting of DNA. Cas9 interrogates DNA and binds reversibly to protospacer adjacent motif (PAM) sequences with stabilization of Cas9 at the target occurring via formation of the *trans*-activating crRNA (tracrRNA)::CRISPR RNA (crRNA) duplex. Mature crRNA anneals to the target with base-pairing complementarity and tracrRNA functional modules govern Cas9 activity and define orthogonality. Activation of Cas9 causes simultaneous cleavage of each strand by the RuvC and HNH domains, represented here as black wedges.

compounded by the protospacer adjacent motif (PAM), a short conserved sequence that must be proximate to the target protospacer (Figure 1) [14–16]. Cas9 introduces a lethal double-stranded DNA break (DSB) at the target locus, effectively acting as a selection against wild type sequences during genome editing [11,12]. Pre-existing mutations in the population can be selected for or against (Figure 2A) but mutations may be introduced subsequent to targeting by host-repair mechanisms (Figure 2B). Mutations elicited by CRISPR–Cas systems are therefore DNA damage and repair machinery dependent. Heterologous expression of Cas9::sgRNA combinations from *Streptococcus pyogenes* has facilitated high-throughput functional genomics in a multitude of eukaryotic organisms and cell lines [17–19]. CRISPR–Cas-derived genome editing tools have revolutionized genetic and biological research in model eukaryotic organisms on account of their efficiency, affordability, and accessibility.

Despite early proof of concept, only three studies have implemented CRISPR–Cas-mediated genome editing in bacteria [20–22], making this a relatively unexplored and underrepresented application of CRISPR–Cas systems. Nevertheless, genome editing via CRISPR–Cas constitutes a next-generation method for programmable and high-throughput functional genomics in prokaryotic backgrounds. Collectively, these studies substantiate the use of CRISPR–Cas systems as genetic tools in bacteria and contribute to our understanding the fundamental genetic and phenotypic outcomes of targeting bacterial genomes. This review summarizes insights from these foundational experiments, highlights considerations for tool development, identifies potential biological hurdles, and predicts future applications of the technology.

Lethality of targeting genomes

The lethality of Cas-mediated DNA cleavage was first observed in its natural ecological role of targeting bacteriophages and plasmids [1,11,12,23], but self-targeting events are an evolutionary cost of housing active CRISPR–Cas systems. The observation of self-complementary spacers, at one time constituting up to 22% of known spacer targets in lactic acid bacteria [24], emphasizes the potential selective pressure of self-targeting events. Identification of self-complementary spacer targets reveals mutations at those chromosomal loci, suggesting that self-targeting events drive mutation or fixation of pre-existing mutations (Figure 2A). Investigation of spacer acquisition in *Streptococcus thermophilus* during exposure to phage led to the infrequent observation of chromosomal acquisition events, which correlated with the disappearance of clones containing self-targeting spacers [25]. Moreover, several studies in diverse backgrounds have reported the lethality of DNA damage induced by self-targeting CRISPR–Cas systems [20,22,26–31]. Transformation of plasmids eliciting self-targeting by Cas proteins is cytotoxic as measured by the relative reduction in viable transformants recovered compared with transformation of non-self-targeting plasmids [28]. CRISPR-mediated depletion of microbial populations results in 3–5-log reductions in populations exhibiting the target sequence, sometimes approaching the transformation efficiency of the respective bacterial background [20,22,26–31]. DNA cleavage by Cas proteins constitutes a significant threat to the survival and fitness of microorganisms, as demonstrated by growth inhibition and aberrant cellular morphology phenotypes consistent with DNA damage observed in *Pectobacterium atrosepticum* following self-targeting events [20]. Active CRISPR–Cas systems cannot coexist in the same cell as the target DNA, which compounds the pressure for mutations to occur, as restoration of the target locus to the wild type does not circumvent CRISPR targeting [32]. Consequently, high- fidelity repair mechanisms are not sufficient for the survival of self-targeting events. Thus, targeting by CRISPR–Cas systems is a selection against the cell populations exhibiting the target genotype. Selection for pre-existing mutations in genetically heterogeneous cell populations supports CRISPR–Cas-directed genome evolution at the population level (Figure 2) [20]. This phenomenon was demonstrated experimentally by the transformation of strain-specific self-targeting plasmids into heterogeneous populations comprising highly similar *Escherichia coli* strains [29], in which dose-dependent depletion of specific population subsets was achieved. This study also reported that lethality was independent of chromosome location, the transcriptional activity of the target, strand bias, and coding versus noncoding regions. Collectively, the well-established lethality of self-targeting events substantiates the utility of CRISPR–Cas for mutagenesis in bacterial genomes by selecting for non-wild type clonal variants. Moreover, they highlight the potential application of Cas-cleavage-driven, sequence-specific evolution of bacterial genomes in mixed populations.

CRISPR–Cas targeting escape strategies

Bacterial cells containing target DNA sequences are efficiently cleared from the population, partially due to the

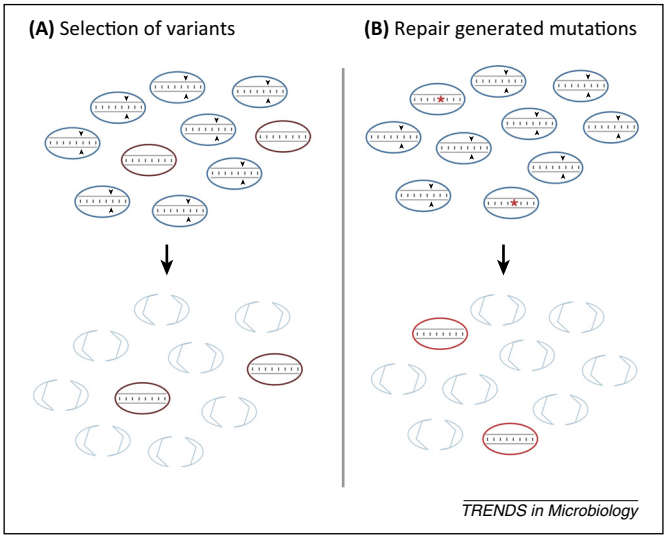

Figure 2. Manipulation of microbial composition in defined consortia. **(A)** A heterogeneous population comprising wild type (turquoise) and clones with pre-existing mutations at the target locus (maroon). Clustered regularly interspaced short palindromic repeats (CRISPR)–CRISPR-associated (Cas) selection against the wild type is applied, driving both the change in microbial composition and population genetics. **(B)** Application of CRISPR–Cas selective pressure against the wild type may result in mutations (denoted by asterisks) through host-dependent repair subsequent to cleavage.

low capacity of bacterial DNA repair mechanisms (Box 1) to cope with Cas cleavage. However, some clones are able to maintain wild type target sequences in the presence of CRISPR–Cas targeting. Genetic analysis of transformants recovered following self-targeting revealed that bacteria can escape targeting by mutation/deletion of the plasmid-encoded spacer or chromosomally encoded Cas machinery to effectively preclude self-targeting (Figure 3A) [20,22,33]. Mutation of the PAM sequence is a major mechanism by which CRISPR–Cas targeting may be circumvented (Figure 3B) [15,20,21,33,34]. The seed sequence comprises the 8–12 bp most proximate to the PAM and is fundamentally involved in the hybridization of crRNA to the cognate target DNA, such that mutations in the seed also abolish targeting (Figure 3B) [21,35,36]. Point mutations within the protospacer sequence are relatively well tolerated and generally do not prevent targeting [20,21]. This is especially true for the PAM-distal spacer sequence, which is consistent with the removal of 10 or 11 bp from the 5′ end of the spacer during the maturation of crRNAs [37]. Instead, deletion of the protospacer constitutes a means of evading Cas-mediated cleavage (Figure 3B). Predictably, mutations that interfere with the biogenesis of crRNAs or the activity of tracrRNA may also abolish activity, but inactivation of CRISPR–Cas systems constitutes a significant cost to the cell. Bacterial means of escape mirror those of the predominant target of CRISPR–Cas systems, phage populations. In phages, mechanisms facilitating circumvention of CRISPR–Cas targeting typically involve alteration of the PAM, seed, or protospacer sequences, achieved through recombination or spontaneous mutation [14,25,34]. Phage populations are inherently genetically diverse yet exhibit a high frequency of homologous recombination (HR) events in response to CRISPR–Cas targeting [25]. Recombination in bacterial populations is limited by natural barriers preventing access to homologous yet variable DNA segments. However, Jiang and coworkers [22] demonstrated that introduction of exogenous genomic DNA (gDNA) caused recombination-mediated survival in *Streptococcus pneumoniae*, which suggests potential roles for competence and horizontal gene transfer in survival for self-targeting events in bacteria.

Box 1. DNA repair mechanisms in bacteria

HR

DNA repair pathways may result in extensive mutation to restore both single-stranded DNA and dsDNA damage caused by Cas-driven cleavage. HR is the most universal and well-characterized DNA repair mechanism in bacteria (Figure 4A) [55]. RecBCD and RecF proteins conduct two HR-mediated DSB repair pathways in *Escherichia coli*. Induction of a DSB elicits helicase activity from RecD and RecB, of which the latter also acts as an exonuclease. Recognition of a chi site slows the complex and causes conformational changes [56], facilitating the loading of RecA proteins onto the residual single-stranded overhang. The RecA–DNA complex then undergoes conformational proofreading and eventual strand invasion of a homologous DNA segment. Resolution of the DNA heteroduplex results in either reciprocal or nonreciprocal recombination. The RecF pathway exhibits partial redundancy for the RecBCD capacity to repair DSBs in bacteria [57]. In this pathway, the RecQ helicase unwinds the blunt-ended DNA and RecJ nuclease processes a 3′ overhang allowing RecFOR to recruit RecA to the strand. The remaining stages of strand invasion and resolution are similar to those for the RecBCD pathway.

A-EJ

A-EJ is a DSB repair mechanism dependent on RecBCD exonuclease end resection and microhomology-mediated recombination (Figure 4B) [58]. RecBCD processivity is thought to potentiate homology searching and leads to the characteristic outcome of variable deletions at the DSB site. In A-EJ, blunt-ended DNA is resected until the 1–9-nt sites of homologous sequences overlap, facilitating LigA-mediated resolution of chromosomal nicks. A-EJ also affords the potential for the integration of exogenous and/or unrelated sequence fragments based on microhomologous (1–9 nt) sequences. Due to the partial redundancy in the repair of DSBs and the subtlety of the mechanism, the characterization of this pathway is not trivial and thus far its distribution is unknown beyond its function in *E. coli*. Thus, it remains difficult to predict the potential for A-EJ repair to occur without prior characterization of this repair pathway in the background microorganisms of interest.

NHEJ

NHEJ occurs through the blunt-ended DNA binding protein Ku and a novel ligase, LigD (Figure 4D) [59]. End resection of blunt DNA results in variable deletions and insertions during repair. A survey of bacterial genome sequences performed by Aravind and Koonin [60] revealed a few bacteria harboring Ku and ligase IV homologs, including *Bacillus subtilis*. Ku and LigD in bacteria exhibit some fundamental differences from their eukaryotic counterparts in terms of gene architecture and mechanism of action [60]. Specifically, Ku homologs exist as genetically fused heterodimers in bacteria in an operon with LigD, although certain systems in *Mycobacterium* spp. encode redundant ligases [59]. To date, NHEJ has been best characterized in *B. subtilis* [61], *Pseudomonas aeruginosa* [62], and *Mycobacterium* spp. [63], but no study has yet investigated the utility of CRISPR–Cas self-targeting in tandem with NHEJ for bacterial genome editing.

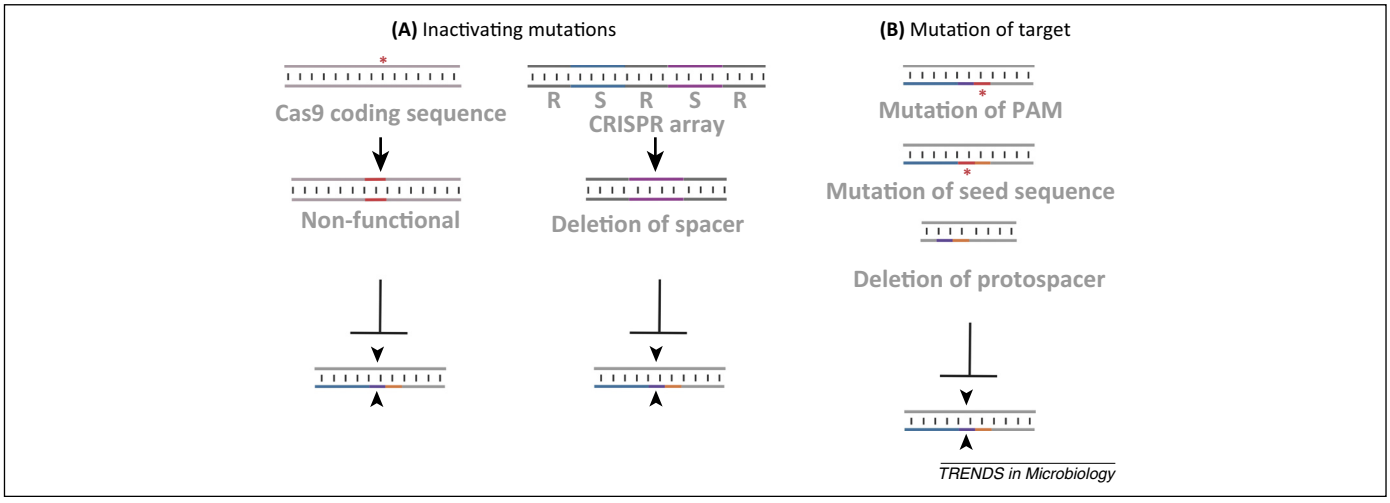

Figure 3. Bacterial mechanisms of escape from clustered regularly interspaced short palindromic repeats (CRISPR)–CRISPR-associated (Cas) targeting. **(A)** Inactivation of Cas genes through mutation prevents targeting but comes at the cost of losing the function of the system. By contrast, recombination between repeats in the CRISPR locus facilitates loss of self-complementary spacers. **(B)** Mutations (indicated by asterisks) that prevent recognition and cleavage by the Cas ribonucleoprotein complex. These mutations include alterations of the protospacer adjacent motif (PAM) and seed sequences or deletion of the protospacer.

Lessons from bacterial genome editing studies

The lethal effects of CRISPR–Cas self-targeting in bacteria are well reported, but few studies have investigated the molecular outcomes of self-targeting events. Recent work by separate groups has provided invaluable insights into the intersection of genome homeostasis and CRISPR–Cas self-targeting [20–22]. These foundational studies pave the way for widespread implementation of CRISPR–Cas technology as a genome editing tool in bacteria. The experiments revealed the mechanistic underpinnings of CRISPR–Cas targeting, exploited DNA repair/replication pathways for designed genome edits, and delineated the genomic plasticity of bacterial populations using CRISPR–Cas targeting.

Homology directed repair in S. pneumoniae

Jiang *et al.* [22] was a landmark study of genome editing using CRISPR–Cas9 in general and was also the first to demonstrate bacterial genome editing. The study determined that double-crossover HR with a donor template restored CRISPR–Cas9-effected chromosomal injury in *S. pneumoniae*. A prophage served as the target for Cas9 cleavage using two derivative strains of *S. pneumoniae* differing from the wild type in an integrated prophage at the *srtA* locus and a prophage integrated strain with a mutated PAM site. Transformation of *S. pneumoniae* cr6 gDNA encoding the prophage targeting Cas9::sgRNA was expected to be lethal to the prophage-harboring strain but not for the strain with the mutated PAM sequence. However, HR at the *srtA* locus was observed, ultimately resulting in the efficient recovery of recombinant clones with deleted prophage genotypes. A similar result was achieved on cotransformation of a wild type *srtA* linear editing template with the prophage targeting Cas9::sgRNA (Figure 4A). The study performed thorough assessments of protospacer and PAM mutations that circumvent Cas9 targeting, providing characterization of the 5′-NGG-3′ PAM requirements and seed sequence of the Type II system from *S. pyogenes*. Moreover, the study highlighted the ability to introduce targeted missense mutations as well as whole-gene deletions using double-crossover HR in the β-galactosidase-encoding gene in *S. pneumoniae*. To assess the efficiency of bacterial genome editing with or without the assistance of Cas9 cleavage, the authors quantified the mutation rate of an artificial stop codon in an erythromycin resistance gene with the outcome restoring the Em^R phenotype. The experiment revealed marginal induction of recombination through Cas9 targeting of the stop codon, but even in the absence of Cas9 cleavage a subpopulation of cells appeared to undergo transformation or recombination at higher frequencies. This study not only established the utility of Cas9-mediated genome editing, but elucidated the molecular underpinnings of the efficiency and limitations of the system.

Efficient and targeted mutagenesis with recombineering in Lactobacillus reuteri

Oh and van Pijkeren employed CRISPR–Cas9 self-targeting in tandem with recombineering for the selection of desired mutations to achieve targeted mutagenesis at nucleotide resolution in *L. reuteri* (Figure 4C) [21]. Plasmid-based expression of *recT*, a single-stranded DNA-binding protein, and *cas9* was used for single-step and dual-step strategies by introducing single-stranded oligonucleotides conferring circumvention of Cas9 targeting. The oligonucleotides were designed to harbor a non-targeted PAM sequence to effectively circumvent Cas9 cleavage and avoid competitive Cas9 binding of the oligonucleotide and possible displacement of the RecT protein [38].This study demonstrated that low-frequency mutations such as whole-gene deletions could be introduced in the cell populations using single-stranded recombineering and that these mutations could be selected for by applying CRISPR–Cas targeting against the wild type genotype. The study also demonstrates the robustness of the system for the mutational biochemical characterization of proteins through introducing missense mutations using Cas9-assisted codon saturation. Similarly to Jiang *et al.* [22], the authors address concerns of making mutations

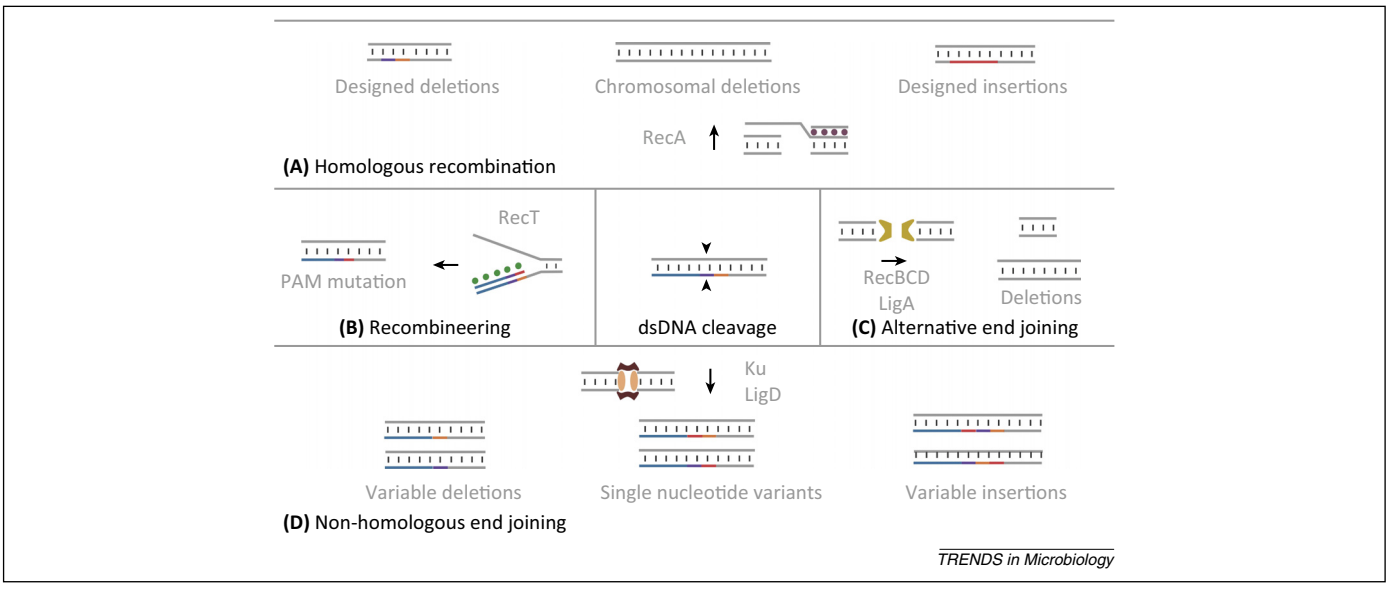

Figure 4. The 'compass rose' of bacterial genome editing via clustered regularly interspaced short palindromic repeats (CRISPR)–CRISPR-associated (Cas) systems and mechanisms of DNA repair in bacteria. In the context of CRISPR–Cas-enhanced genome editing, recombination **(A)** can occur with a non-wild type editing template or between homologous sequences flanking the target sequence. **(B)** Recombineering strategy for genome editing is achieved through heterologous expression of Cas9 and recT followed by transformation of a single-stranded oligonucleotide with identity to the target sequence. The single-stranded oligonucleotide anneals to the lagging strand as a primer during discontinuous DNA replication, thus incorporating the desired mutations. The alternative pathway is mediated by the RecBCD complex and LigA **(C)**, whereas the classical nonhomologous end joining pathway **(D)** is conducted by Ku and LigD. Both result in variable deletions and insertions.

at additional non-PAM sites while still using the highly effective PAM-mutation-based circumvention. As reported by Oh and van Pijkeren, recombineering represents a precise and efficacious genetic tool in *L. reuteri*, especially when applying CRISPR–Cas-assisted selection for desired mutations. However, successful application of this technology may require considerable optimization for use in disparate backgrounds and the universal efficacy of the method remains undetermined given the significant variation in the capacity of bacteria to perform recombineering.

Alternative end joining (A-EJ) and large deletions in Pectobacterium atrosepticum

In contrast to the other bacterial genome editing studies, Vercoe *et al.* [20] used a natively active Type I-F system, which is less suited for introducing designed mutations due to the unpredictable and extensive nature of the DNA damage caused by Cas3 exonuclease activity. It was demonstrated that lethality due to a self-targeting spacer in *P. atrosepticum* was abrogated by a single PAM sequence mutation, but self-targeting was restored through plasmid-based expression of a programmable repeat–spacer array. In the absence of a donor template for HR, large deletions were achieved through recombination of mobile genetic element features (Figure 4A). Deletion of the entire pathogenicity island ($\sim$98 kb) occurred reproducibly through the recombination of *attL* and *attR* sites flanking the island. Interestingly, the authors reported that the deletion occurred spontaneously at low frequency in wild type cell populations as detected by PCR amplification of the *attB* excision footprint. By contrast, variable deletions within the pathogenicity island consistent with A-EJ also occurred through microhomologous sequences. The two mechanisms of mutation observed in this study highlight the ordinal effect of mutation and DNA damage, since it can be postulated that deletion of the island was observed through selection of pre-existing mutations (Figure 2A), whereas deletions due to A-EJ were likely to have occurred subsequent to cleavage of the chromosome (Figure 2B).

Design considerations

The exceptional range of microbial diversity poses shared challenges for many genetic manipulation strategies. The mechanisms governing DNA homeostasis are highly background specific; thus, genetic tool development is limited by factors including transformation efficiency, plasmid replication, capacity for recombination/integration, DNA methylation, and DNA repair pathways. Similarly, the application of CRISPR–Cas9 technology for bacterial genome editing hinges on these molecular processes, but there are also specific design considerations for accurate and efficient use of the technology.

Design of spacer sequences

Off-target cleavage is expected to occur infrequently in bacterial systems relative to eukaryotes, which can be attributed to the lower occurrence of sequences homologous to a given spacer–PAM combination in smaller genomes. Given that single self-targeting events result in a significant reduction in the recovery of viable transformants, any off-target cleavage leading to multiple events of CRISPR–Cas-induced chromosomal injury in a single cell should compound this reduction in recovery, leading to a decreased incidence of mutation at extraneous loci. Proper selection of spacer sequences is essential to further prevent unintended cleavage events and to maximize efficiency. To date, two strict criteria for the selection and design of spacers are the location of consensus PAM sequences and avoiding incidental sequence identity to

extraneous genomic loci. Putative protospacers are constrained by defining the location of putative PAM sequences in the target locus. Since PAM and seed sequences are integral for recognition and activity at the target, spacers must be selected based on uniqueness of these components to prevent off-target cleavage. In Type II-A systems, approximately 10 nt is removed from the 5′ end of the spacer during crRNA maturation, suggesting that they are irrelevant for target specificity [37]. Protospacers containing sequences identical to the PAM should not be considered, to prevent competitive Cas recruitment that may limit the cleavage of the desired locus [38]. There is no web tool dedicated to spacer design for bacterial genome editing and the utility of current eukaryotic tools for designing bacterial spacers is undetermined.

Increasing transformation efficiency for optimized system delivery

Transformation efficiencies are limiting in many backgrounds, but transformation (natural or induced), transduction, and conjugation are all potential avenues for experiments requiring simultaneous cotransformation of editing templates and expression vectors. The lethality of self-targeting compounds the need for high transformation efficiency of CRISPR–Cas components when interference is the direct result of transformation. Limited transformation efficiency can be compensated for by designing high-frequency-mutation strategies facilitating circumvention of CRISPR–Cas targeting, increasing the recovery of desired genotypes. Development of tightly regulated inducible expression systems bypasses the low transformation efficiency of self-targeting plasmids, allowing the induction of self-targeting in highly concentrated cultures. Inducible systems could therefore increase the recovery of desired mutations and the identification of low-frequency mutations. An intriguing observation noted by Jiang *et al.* [22] suggested that certain bacterial subpopulations were more prone to recombination/transformation. Experiments with multiple rounds of recombination/transformation may therefore constitute short-term directed 'evolution'. This process results in disproportionate selection of the population with a higher competency phenotype, with potential applications in molecular biology for backgrounds with low transformation or recombination efficiencies.

DNA repair mechanisms in bacteria

Since the mutations elicited by CRISPR–Cas systems are both DNA damage and repair machinery dependent, it is prudent to consider the DNA repair pathways present in each microbiological background. The universality of HR and its well-characterized mechanism and relatively high frequency have led to the widespread development and use of HR-mediated genetic technologies in diverse bacterial backgrounds. Therefore, HR can similarly be considered a practical and applicable repair pathway for the introduction of mutations in CRISPR–Cas-assisted genome editing. HR-based repair of DSBs and the simultaneous generation of desired mutations can be achieved through the provision of a homologous editing template, which affords the potential for both deletions and gene replacements. Recombination can occur between native homologous sequences flanking DSBs in the genome, resulting in the deletion of large genomic segments [20]. To this end, native repair pathways can be exploited for the generation of desired mutations or enzymatic machinery can be heterologously expressed to introduce repair pathway platforms for mutagenesis. In particular, recombinant expression of proteins Ku and LigD in bacteria may offer an avenue for the high-frequency generation and recovery of mutants. By contrast, native pathways may introduce undesired mutations at high frequencies relative to that of designed genetic outcomes, making it necessary to knock out or transcriptionally downregulate certain pathways to prevent undesired repair from occurring. Conversely, targeting genomes with CRISPR–Cas systems also affords the potential for the characterization of native DNA repair pathways in diverse microbial backgrounds.

Future applications

CRISPR–Cas selection against target sequences has already been tangibly exploited in a few bacteria to elicit genome edits, but further development of the technology has considerable potential for revolutionizing bacterial genetics and genomics.

Cas9 and editing template delivery through transformation of linear nucleic acids

Linear DNA transformation has not been broadly applied as a tool for genetic manipulation in bacteria, since replication of DNA requires circularity and host exonuclease activity causes rapid degradation of linear dsDNA. A few studies have employed linear DNA to elicit gene replacement by double-crossover recombination, with success in *E. coli*, *Bordetella pertussis*, *S. thermophilus*, and *S. pneumoniae* [22,39–41]. Synthetic DNA molecules are easily designed and affordable for engineered mutagenesis and splicing by overlap extension PCR also allows the generation of editing templates [42]. Thus, linear DNA transformation techniques may increasingly be considered a viable option for genome editing when complemented with CRISPR–Cas selection against the wild type. Transformation of linear RNA is another unexplored avenue for the mutagenesis of bacteria, but with the strong selective pressure of Cas9 targeting, RNA-based expression of Cas9 and sgRNAs holds potential for increased throughput in genome editing of bacteria.

Exploitation of endogenous and orthogonal systems

CRISPR–Cas systems are found in approximately 46% of bacteria and 84% of Archaea [43], highlighting the potential for genome editing applications using endogenously active systems [44]. However, there is a general paucity of systems with characterized PAMs and confirmed activity in interference [45]. Self-targeting or plasmid interference assays must first be performed to ensure their activity and assess PAM recognition. Platforms for genome editing in CRISPR-deficient backgrounds may be generated through plasmid or chromosomal expression of vested Cas9s and sgRNAs [46]. The use of multiple orthogonal systems offers the advantage of utilizing disparate PAM sequences,

increasing the range of target sequences without sacrificing specificity [47]. Longer PAMs may offer increased specificity and potentially decreased off-target cleavage. Extended PAMs may increase the efficacy of cleavage by further compounding the low frequency at which point mutations effectively confer circumvention of Cas9 recognition. It is noteworthy, however, that recent evidence indicates that biochemical recognition of PAMs is not as stringent as bioinformatically determined consensus sequences would suggest [48]. Therefore, it may be necessary to empirically determine the specific nucleotides contributing to stringent PAM recognition for each CRISPR–Cas system.

Understanding bacterial genome biology

The high prevalence of mobile genetic elements in bacterial genomes [49] presents a unique challenge for eliciting targeted mutations at these loci. Excision of mobile genetic elements may occur in the face of CRISPR–Cas selective pressure [20], dependent on the frequency of excision relative to that of the desired mutation. Therefore, *in silico* prediction of mobile genomic segments may be used to identify potentially expendable regions, which can then be experimentally validated with CRISPR–Cas selection. Despite the potential difficulty of generating designed mutations in these segments, there are considerable applications for the excision of mobile genetic elements, such as defining minimal bacterial genomes and the characterization of putative unannotated proteins and essential genes. CRISPR–Cas selection can therefore be used to screen for clonal subtypes within heterogeneous populations, delineating the locus-dependent plasticity of bacterial genomes.

Concluding remarks

Development of CRISPR–Cas technology in bacteria has yielded applications in typing and strain detection [50,51], the exploitation of natural/engineered immunity against mobile genetic elements [52–54], the manipulation of microbial consortia/generation of smart antibiotics [29], and programmable transcriptional regulation [45]. However, few studies have focused on the development of CRISPR–Cas genome editing tools in bacteria. This streamlined methodology holds potential for increasing expediency and efficiency in the generation of desired mutations, potentially without the necessity of plasmid integration, extensive screening, or counter-selection. Microbial diversity necessitates the development of efficient transformation protocols and genetic tools for bacterial genome editing (Box 2), but CRISPR–Cas technology opens new avenues in genetic engineering applications.

Box 2. Outstanding questions

- What are the most practical ways of combining traditional genetic strategies with CRISPR–Cas selection?
- What proportion of endogenous and active systems can be repurposed for genome editing or gene regulation?
- Can NHEJ be broadly exploited for genome editing in bacteria?
- How can the development of orthogonal systems into genome editing tools expand the repertoire for genetic manipulation in bacteria?

Acknowledgments

The authors thank all members of the CRISPR and TRK laboratory for constructive insights, especially Brant Johnson for reviewing the manuscript. This work was supported by start-up funds from North Carolina State University, a grant from the National Science Foundation (CBET-1403135), and the North Carolina Agriculture Foundation.

Disclaimer statement

K.S. and R.B. are inventors on several patents related to CRISPR–Cas systems and their various uses. R.B. is on the board of directors of Caribou Biosciences and cofounder of Intellia Therapeutics.

References

1 Barrangou, R. *et al.* (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* 315, 1709–1712
2 Barrangou, R. and Marraffini, L.A. (2014) CRISPR–Cas systems: prokaryotes upgrade to adaptive immunity. *Mol. Cell* 54, 234–244
3 Brouns, S.J.J. *et al.* (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* 321, 960–964
4 Young, J.C. *et al.* (2012) Phage-induced expression of CRISPR-associated proteins is revealed by shotgun proteomics in *Streptococcus thermophilus*. *PLoS ONE* 7, e38077
5 Goh, Y.J. *et al.* (2011) Specialized adaptation of a lactic acid bacterium to the milk environment: the comparative genomics of *Streptococcus thermophilus* LMD-9. *Microb. Cell Fact.* 10 (Suppl. 1), S22
6 Makarova, K.S. *et al.* (2011) Evolution and classification of the CRISPR–Cas systems. *Nat. Rev. Microbiol.* 9, 467–477
7 Sinkunas, T. *et al.* (2011) Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J.* 30, 1335–1342
8 Deltcheva, E. *et al.* (2011) CRISPR RNA maturation by *trans*-encoded small RNA and host factor RNase III. *Nature* 471, 602–607
9 Karvelis, T. *et al.* (2013) crRNA and tracrRNA guide Cas9-mediated DNA interference in *Streptococcus thermophilus*. *RNA Biol.* 10, 841–851
10 Hale, C.R. *et al.* (2009) RNA-guided RNA cleavage by a CRISPR RNA–Cas protein complex. *Cell* 139, 945–956
11 Garneau, J.E. *et al.* (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* 468, 67–71
12 Gasiunas, G. *et al.* (2012) Cas9–crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U.S.A.* 109, E2579–E2586
13 Jinek, M. *et al.* (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* 337, 816–821
14 Deveau, H. *et al.* (2008) Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* 190, 1390–1400
15 Horvath, P. *et al.* (2008) Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J. Bacteriol.* 190, 1401–1412
16 Mojica, F.J.M. *et al.* (2009) Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* 155, 733–740
17 Mali, P. *et al.* (2013) RNA-guided human genome engineering via Cas9. *Science* 339, 823–826
18 Hsu, P.D. *et al.* (2014) Development and applications of CRISPR–Cas9 for genome engineering. *Cell* 157, 1262–1278
19 Sander, J.D. and Joung, J.K. (2014) CRISPR–Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* 32, 347–355
20 Vercoe, R.B. *et al.* (2013) Cytotoxic chromosomal targeting by CRISPR/Cas systems can reshape bacterial genomes and expel or remodel pathogenicity islands. *PLoS Genet.* 9, e1003454
21 Oh, J-H. and van Pijkeren, J-P. (2015) CRISPR–Cas9-assisted recombineering in *Lactobacillus reuteri*. *Nucleic Acids Res.* 42, e131
22 Jiang, W. *et al.* (2013) RNA-guided editing of bacterial genomes using CRISPR–Cas systems. *Nat. Biotechnol.* 31, 233–239
23 Levin, B.R. *et al.* (2013) The population and evolutionary dynamics of phage and bacteria with CRISPR-mediated immunity. *PLoS Genet.* 9, e1003312
24 Horvath, P. *et al.* (2009) Comparative analysis of CRISPR loci in lactic acid bacteria genomes. *Int. J. Food Microbiol.* 131, 62–70
25 Paez-Espino, D. *et al.* (2013) Strong bias in the bacterial CRISPR elements that confer immunity to phage. *Nat. Commun.* 4, 1430
26 Edgar, R. and Qimron, U. (2010) The *Escherichia coli* CRISPR system protects from λ lysogenization, lysogens, and prophage induction. *J. Bacteriol.* 192, 6291–6294

27 Marraffini, L.A. and Sontheimer, E.J. (2008) CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* 322, 1843–1845
28 Bikard, D. *et al.* (2012) CRISPR interference can prevent natural transformation and virulence acquisition during *in vivo* bacterial infection. *Cell Host Microbe* 12, 177–186
29 Gomaa, A.A. *et al.* (2014) Programmable removal of bacterial strains by use of genome-targeting CRISPR–Cas systems. *MBio* 5, e00928–e1013
30 Bikard, D. *et al.* (2014) Exploiting CRISPR–Cas nucleases to produce sequence-specific antimicrobials. *Nat. Biotechnol.* 32, 1146–1150
31 Citorik, R.J. *et al.* (2014) Sequence-specific antimicrobials using efficiently delivered RNA-guided nucleases. *Nat. Biotechnol.* 32, 1141–1145
32 Jiang, W. *et al.* (2013) Dealing with the evolutionary downside of CRISPR immunity: bacteria and beneficial plasmids. *PLoS Genet.* 9, e1003844
33 Stern, A. *et al.* (2010) Self-targeting by CRISPR: gene regulation or autoimmunity? *Trends Genet.* 26, 335–340
34 Sun, C.L. *et al.* (2013) Phage mutations in response to CRISPR diversification in a bacterial population. *Environ. Microbiol.* 15, 463–470
35 Semenova, E. *et al.* (2011) Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10098–10103
36 Wiedenheft, B. *et al.* (2011) RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10092–10097
37 Carte, J. *et al.* (2014) The three major types of CRISPR–Cas systems function independently in CRISPR RNA biogenesis in *Streptococcus thermophilus*. *Mol. Microbiol.* 93, 98–112
38 Sternberg, S.H. *et al.* (2014) DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* 507, 62–67
39 Jasin, M. and Schimmel, P. (1984) Deletion of an essential gene in *Escherichia coli* by site-specific recombination with linear DNA fragments. *J. Bacteriol.* 159, 783–786
40 Zealey, G.R. *et al.* (1990) Gene replacement in *Bordetella pertussis* by transformation with linear DNA. *Nat. Biotechnol.* 8, 1025–1029
41 Fontaine, L. *et al.* (2010) A novel pheromone quorum-sensing system controls the development of natural competence in *Streptococcus thermophilus* and *Streptococcus salivarius*. *J. Bacteriol.* 192, 1444–1454
42 Horton, R.M. *et al.* (1989) Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* 77, 61–68
43 Grissa, I. *et al.* (2007) CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 35, W52–W57
44 Luo, M.L. *et al.* (2014) Repurposing endogenous Type I CRISPR–Cas systems for programmable gene repression. *Nucleic Acids Res.* 43, 674–681
45 Bondy-Denomy, J. and Davidson, A.R. (2014) To acquire or resist: the complex biological effects of CRISPR–Cas systems. *Trends Microbiol.* 22, 218–225
46 Sapranauskas, R. *et al.* (2011) The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* 39, 9275–9282
47 Briner, A.E. *et al.* (2014) Guide RNA functional modules direct Cas9 activity and orthogonality. *Mol. Cell* 56, 333–339
48 Esvelt, K.M. *et al.* (2013) Orthogonal Cas9 proteins for RNA-guided gene regulation and editing. *Nat. Methods* 10, 1116–1121
49 Darmon, E. and Leach, D.R.F. (2014) Bacterial genome instability. *Microbiol. Mol. Biol. Rev.* 78, 1–39
50 Groenen, P.M. *et al.* (1993) Nature of DNA polymorphism in the direct repeat cluster of *Mycobacterium tuberculosis*; application for strain differentiation by a novel typing method. *Mol. Microbiol.* 10, 1057–1065
51 Shariat, N. *et al.* (2013) The combination of CRISPR–MVLST and PFGE provides increased discriminatory power for differentiating human clinical isolates of *Salmonella enterica* subsp. *enterica* serovar Enteritidis. *Food Microbiol.* 34, 164–173
52 Barrangou, R. *et al.* (2013) Genomic impact of CRISPR immunization against bacteriophages. *Biochem. Soc. Trans.* 41, 1383–1391
53 Barrangou, R. and Horvath, P. (2012) CRISPR: new horizons in phage resistance and strain identification. *Annu. Rev. Food Sci. Technol.* 3, 143–162
54 Hynes, A.P. *et al.* (2014) Adaptation in bacterial CRISPR–Cas immunity can be driven by defective phages. *Nat. Commun.* 5, 4399
55 Wigley, D.B. (2013) Bacterial DNA repair: recent insights into the mechanism of RecBCD, AddAB and AdnAB. *Nat. Rev. Microbiol.* 11, 9–13
56 Dixon, D.A. and Kowalczykowski, S.C. (1993) The recombination hotspot chi is a regulatory sequence that acts by attenuating the nuclease activity of the *E. coli* RecBCD enzyme. *Cell* 73, 87–96
57 Morimatsu, K. and Kowalczykowski, S.C. (2003) RecFOR proteins load RecA protein onto gapped DNA to accelerate DNA strand exchange: a universal step of recombinational repair. *Mol. Cell* 11, 1337–1347
58 Chayot, R. *et al.* (2010) An end-joining repair mechanism in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 2141–2146
59 Shuman, S. and Glickman, M.S. (2007) Bacterial DNA repair by non-homologous end joining. *Nat. Rev. Microbiol.* 5, 852–861
60 Aravind, L. and Koonin, E.V. (2001) Prokaryotic homologs of the eukaryotic DNA-end-binding protein Ku, novel domains in the Ku protein and prediction of a prokaryotic double-strand break repair system. *Genome Res.* 11, 1365–1374
61 Weller, G.R. *et al.* (2002) Identification of a DNA nonhomologous end-joining complex in bacteria. *Science* 297, 1686–1689
62 Zhu, H. and Shuman, S. (2005) Novel 3′-ribonuclease and 3′-phosphatase activities of the bacterial non-homologous end-joining protein, DNA ligase D. *J. Biol. Chem.* 280, 25973–25981
63 Gong, C. *et al.* (2005) Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C. *Nat. Struct. Mol. Biol.* 12, 304–312