Review

Exploiting CRISPR/Cas: Interference Mechanisms and Applications

Hagen Richter, Lennart Randau and André Plagens *

Prokaryotic Small RNA Biology, Max Planck Institute for Terrestrial Microbiology,
Karl-von-Frisch-Straße 10, 35043 Marburg, Germany;
E-Mails: hagen.richter@mpi-marburg.mpg.de (H.R.); lennart.randau@mpi-marburg.mpg.de (L.R.)

* Author to whom correspondence should be addressed; E-Mail: andre.plagens@mpi-marburg.mpg.de;
Tel.: +49-0-6421-178-650; Fax: +49-0-6421-178-599.

Abstract: The discovery of biological concepts can often provide a framework for the development of novel molecular tools, which can help us to further understand and manipulate life. One recent example is the elucidation of the prokaryotic adaptive immune system, clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) that protects bacteria and archaea against viruses or conjugative plasmids. The immunity is based on small RNA molecules that are incorporated into versatile multi-domain proteins or protein complexes and specifically target viral nucleic acids via base complementarity. CRISPR/Cas interference machines are utilized to develop novel genome editing tools for different organisms. Here, we will review the latest progress in the elucidation and application of prokaryotic CRISPR/Cas systems and discuss possible future approaches to exploit the potential of these interference machineries.

Keywords: CRISPR; crRNA; Cas9; Cascade; interference; genome editing; RGEN; TALEN; ZNF

1. Introduction

The clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) systems are found in many bacteria and nearly all archaea and constitute an adaptive immune system that recognizes and prevents viral attacks. A breakthrough for the basic understanding of the immunity

mechanism was achieved by Barrangou and coworkers, who could show that *Streptococcus thermophilus* can acquire resistance against a bacteriophage by integrating a genome fragment of an infectious virus into its CRISPR locus [1]. A CRISPR cluster is a genomic DNA element that consists of a series of short repeat sequences (typically 24–37 bp) that are separated by unique spacer sequences of similar length [2]. These sequences are often fragments derived from a viral genome, illustrating the genetic memory of previous infections [3,4]. The second part of the CRISPR/Cas machinery is encoded in the Cas genes, and the Cas proteins fulfill essential functions within the immunity mechanism.

Figure 1. The interference step in the three clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) types. Indicated for all three systems are the targeted DNA region (blue), the targeting crRNA (red) and the protospacer adjacent motif (PAM, green). In type I systems (**A**), the invading DNA is recognized by the Cascade:crRNA complex. The PAM motif promotes the identification of the foreign DNA. Subsequently, the nuclease, Cas3, is recruited and degrades the target DNA. Type II systems (**B**) require only Cas9 for interference and do not rely on a multi-protein complex. Complex formation of Cas9, tracrRNA and pre-crRNA enables RNase III to mature crRNAs. The resulting complex of Cas9 and the tracrRNA:crRNA duplex recognizes the invading nucleic acid, and the Cas9 nuclease generates blunt-ended cleavage of both DNA strands. In type III systems (**C**), a multi-protein complex (Csm or Cmr) or Cas6 processes pre-crRNA into mature crRNA. The complex-bound crRNA recognizes invading DNA (Csm) or RNA (Cmr), resulting in target degradation.
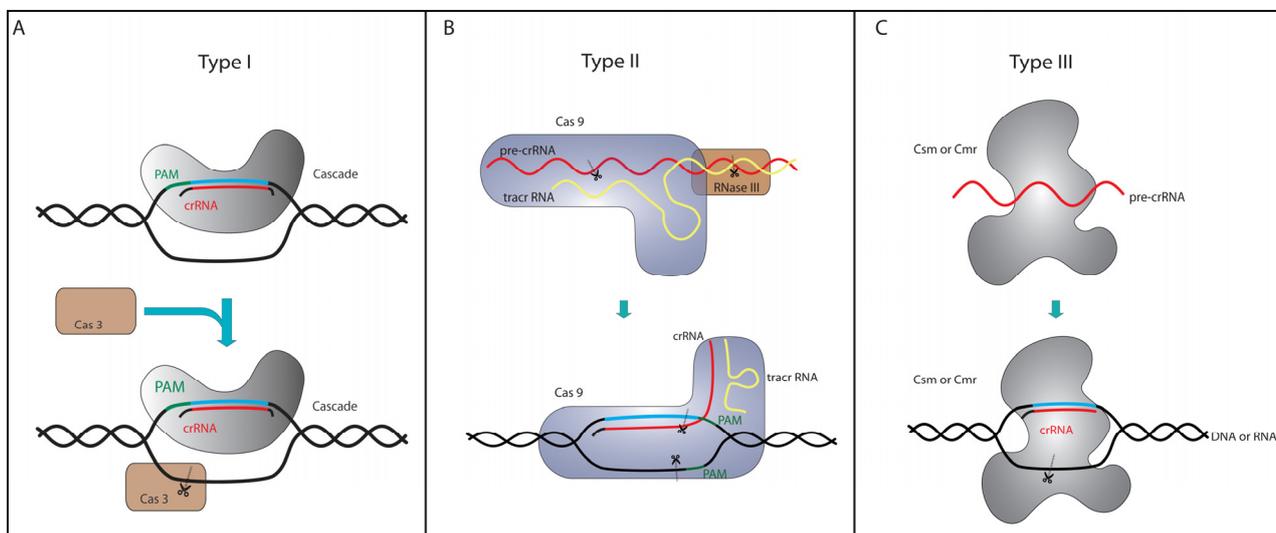
The CRISPR/Cas defense response is classified into three main stages, which are conserved throughout all CRISPR systems. In the first stage, termed adaptation, the injected viral DNA is recognized, and a fragment of this DNA inserted as a new spacer into a host CRISPR array. A short conserved sequence (2 to 5 nt), called the protospacer adjacent motif (PAM), flanks the spacer sequence in the viral genome (termed protospacer) and, thus, determines the targets of most CRISPR/Cas systems [5,6]. A new spacer is always integrated at the AT-rich leader site of a CRISPR, which is proposed to contain specific sequence elements that direct spacer DNA addition [7,8]. The detailed mechanism of spacer acquisition is only partially understood, but the highly conserved

proteins, Cas1 and Cas2, were identified as key players in this process [9,10]. The second stage is characterized by the transcription of a CRISPR cluster into a long precursor-crRNA (pre-crRNA). Transcription is controlled by sequences in the leader region that contains promoter elements and possible binding sites for regulatory proteins [11,12]. In many systems, the long pre-crRNA transcript is processed by a Cas endonuclease (Cas6) into short crRNAs [13–17]. A typical crRNA generated by Cas6 cleavage contains the complete spacer sequence flanked by an 8 nt 5'-hydroxyl repeat tag and a 2'–3' cyclic phosphate repeat end [17,18]. In some cases, 3' termini of mature crRNAs are trimmed further, but the mechanism is unknown. In the final stage, the interference reaction, these mature crRNAs are incorporated into a larger Cas protein complex and used to target and degrade the viral DNA during a repeated attack (Figure 1) [19–21]. In the next section, we will focus on the molecular details of different interference mechanisms with diverse sets of Cas proteins and multi-subunit complexes. These crRNA/Cas protein complexes exhibit a large potential for the development of genetic tools.

2. Three Strategies to Cope with Viruses

CRISPR/Cas systems are identified in diverse bacterial and archaeal species, often living in the most extreme ecological niches. Some basic principles of the CRISPR/Cas systems are conserved within all prokaryotes, but the plethora of identified Cas protein families reflects the divergence of mechanistic details during evolution. The continuous co-evolution of viruses and their hosts led to the emergence of anti-CRISPR measures in viruses [22], which might explain the necessity for CRISPR/Cas diversification. Computational studies classified all CRISPR/Cas systems into three types and at least ten subtypes [23–25].

Type I CRISPR/Cas systems are found in both bacteria and archaea and are comprised of six different subtypes (subtypes I-A to I-F). The essential and significantly conserved marker protein in the interference reaction is Cas3, which contains a HD phosphohydrolase domain and a DExH-like helicase domain [24,25]. Both domains are also found to be encoded separately by two discrete genes. These two domains have been shown to unwind dsDNA (helicase domain) and cleave ssDNA (HD nuclease domain), depending on ATP and Mg^{2+} ions [26–28]. Cas3 interacts with a complex of different Cas proteins that bind and deliver the crRNA (Figure 1A). This complex is termed Cascade (CRISPR-associated complex for antiviral defense) and is best studied for the type I-E system from *Escherichia coli* [18,19]. The type I-E Cascade was shown to bind the short crRNAs, which are utilized to recognize DNA target sequences that are complementary to the spacer sequence within the crRNA. Subsequently, Cascade recruits Cas3 to degrade the targeted viral DNA molecule, relying on negatively supercoiled DNA [29,30]. The first 6–12 nt of the crRNA spacer are most important for target binding and are termed the seed sequence [20,21]. The point mutation of single seed sequence nucleotides results in drastic binding defects. However, mismatches in the crRNA spacer sequence following the seed sequence are tolerated, and the mutated crRNA can still be bound [31]. To avoid that the Cascade:Cas3 complex degrades the host genomic encoded CRISPR cluster, it has to be ensured that the 5' terminal tag of the crRNA and the PAM sequence located upstream of the viral protospacer do not form base pairs. To discriminate between self- and non-self- targets, Cascade screens and specifically binds a PAM sequence, which results in helical destabilization and strand

invasion of the matching seed sequence [32,33]. The PAM sequence for type I systems is typically 2–3 bases long and can differ between different subtypes and even organisms. The I-E Cascade complex has a size of 405 kDa and is composed of the five subunits, Cas6e, Cse1, Cse2, Cas7 and Cas5. Cas7 and Cas5 tightly bind and protect the crRNA from degradation [19], whereas Cse1 and Cse2 were shown to be nucleic acid-binding proteins that preferentially interact with the DNA target [34,35].

The targeting complex facilitates base pairing of the crRNA with the complementary DNA strand and additional displacement of the non-complementary strand to produce a so-called R-loop structure [18,36]. Using cryo-electron microscopy, the overall structure of the I-E Cascade complex was resolved and revealed a general outline that is often described as a seahorse-like shape [18,20]. The conservation of the crRNA-binding subunits, Cas7 and Cas5, as well as Cas3, throughout all type I subtypes, suggest structural and functional similarities of Cascade, but further biochemical and structural data are required to compare complexes and the interference reaction in different prokaryotic families in detail. The related I-A Cascade complex shows a conserved Cas7, Cas5 and crRNA assembly platform, which might recruit Cas3 and additional subtype-specific proteins to form the active interference machinery [36,37]. Plasmid-based interference assays were established that could show *in vivo* interference activity for I-A and I-B Cascade modules [38,39].

Type II systems have only been found in bacterial genomes and are characterized by a distinct minimal set of *cas* genes [24,25]. In these systems, the large multifunctional protein, Cas9, is involved in both the maturation of crRNAs and in the subsequent interference reaction [40]. The processing of crRNAs is dependent on a trans-activating crRNA (tracrRNA) encoded in the vicinity of CRISPR loci and containing a 25 nt long stretch that is complementary to the crRNA repeat sequence (Figure 1B) [41]. The comparison of several tracrRNA molecules did not identify any highly conserved sequence or structure elements other than the anti-repeat sequence [42]. Cas9 facilitates the base pairing of tracrRNA and pre-crRNA, which form a RNA duplex that is then targeted by the host endonuclease, RNase III. Cleavage of this duplex by RNase III generates mature crRNAs with 20 nt spacer-derived 5'-tags and 19–22 nt repeat-derived 3'-tags [41,43]. In the interference step, the cleavage of target dsDNA requires not only crRNA and Cas9, but also the presence of tracrRNA. Cas9 cleaves the DNA strand complementary to the crRNA with a McrA/HNH nuclease domain and the non-complementary strand with a RuvC-like (RNase H fold) domain in the presence of Mg^{2+} ions [43]. The interference against the viral DNA requires a conserved 5 nt-long PAM sequence (NGGNG), located immediately downstream of the protospacer [5,44]. The precise DNA cleavage site was identified 3 nt upstream of the PAM for the complementary strand, whereas the non-complementary DNA strand is cleaved at additional sites within three to eight base pairs upstream of the PAM, producing blunt-ended cleavage products [43,45].

The two known type III systems (type III-A and type III-B) are predominantly found in archaeal genomes [24,25], and interestingly, type III-B systems are only found in combination with one or more other CRISPR subtypes. Type III systems encode the CRISPR-specific endoribonuclease, Cas6, and the subtype-specific Cas10 protein that is very likely involved in target interference (Figure 1C). Similar to Cas3 proteins of type I systems, Cas10 encodes a HD nuclease domain that is proposed to have similar function in target degradation [24,25]. The type III-A system of *Staphylococcus epidermidis* contains five Csm proteins and was shown to target DNA [46]. DNA targeting by this system does not require a specific PAM sequence, but sequences complementary to the 8 nt 5'-tag of

the crRNA are not targeted by this system [47]. In the type III-B system of *Pyrococcus furiosus*, Cas6 is not an integral part of the interference complex after the crRNA processing, but the 8-nt 5' repeat tag serves as an anchor for the assembly of a six protein (Cmr1–Cmr6) ribonucleoprotein interference complex. A similar Cmr complex with seven proteins (Cmr1–Cmr7) was identified for *Sulfolobus solfataricus* and shown to endonucleolytically cleave invading RNA at UA dinucleotides [48]. Targeting of RNA was shown to be PAM-independent for both investigated Cmr complexes. The crystal structure of a Cmr2–Cmr3 complex revealed a conserved RNA binding surface of Cmr3, which is reminiscent of the Cas6 RNA interaction surface and illustrates the RNA binding ability of this Cmr subunit [49,50]. Notably, these two interference complexes differ from all other investigated subtypes, as they specifically target RNA and not DNA [11,51]. However, recently, it could be demonstrated *in vivo* that Cmr proteins can target also plasmid DNA in a PAM-independent manner [52].

3. CRISPR Systems as Genome Editing Tools

The application of the diverse CRISPR/Cas systems as genetic tools offers great potential. Initially, before the discovery of Cas protein functions, the diversity of CRISPR sequences was mainly utilized in a powerful method to rapidly identify closely related bacterial strains (e.g., *Mycobacterium tuberculosis*). This genotyping method is known as spacer oligonucleotide typing (spoligotyping) [53,54]. The investigation of different Cas protein activities introduced them as diverse components for genetic tool development. One example is the pre-crRNA processing enzyme, Cas6f (previously termed Csy4), which was shown to be useful for predictable gene expression. In these approaches, the Cas6f cleavage site sequence of a CRISPR repeat can be fused to a gene of interest. Processing of this site by Cas6f can then be used to physically separate genetic elements (e.g., untranslated regions, regulatory elements, ribosome binding sites) at the mRNA level. Therefore, transcript impedances (e.g., secondary structures) can be reduced, which results in a more predictable gene expression pattern [55]. Additionally, the high substrate affinity of Cas6f was used to create a specific RNA-binding bait protein. This construct can be used in high-throughput RNA affinity purification protocols to isolate RNA molecules with a 16 nt hairpin sequence derived from the 3'-terminal crRNA repeat tag [56].

The recent analysis of Cas protein interference complexes immediately revealed their great potential for the development of genetic tools that are required to provide specific DNA or RNA targeting. One key player in this development is the large type II protein, Cas9. Qi and colleagues could show that a nuclease inactive mutant of Cas9 in combination with a sequence specific crRNA can be utilized for targeted DNA recognition to interfere with transcriptional elongation, RNA polymerase or transcription factor binding. This gene silencing activity was termed CRISPRi for CRISPR interference in reference to RNAi [57]. Subsequently, other groups reported the utilization of Cas9/crRNA complexes for genome editing in different organisms, e.g., human cell lines, zebrafish, mice, drosophila, yeast and bacteria. Cas9 interference experiments indicated that the fusion product of crRNA and tracrRNA has similar efficiency as the RNase III-processed crRNA:tracrRNA duplex [43]. Therefore, these genome editing approaches use Cas9 together with a specific small guide RNA (sgRNA), which was designed to resemble the fused crRNA/tracrRNA sequence. The resulting ribonucleoproteins are termed RNA guided endonucleases (RGENs) and were shown to target single genes or even multiple genes, allowing efficient and site-specific editing of the target sequence [58–67]. Specificity of the targeting reaction is

determined by the sgRNA sequence, which immediately displays the advantage of this method over other established genome editing methods, like zinc-finger nucleases (ZNF) or transcription activator-like effector nucleases (TALEN) [68]. A single Cas9 protein can be retargeted using different sgRNA sequences and without the time-consuming protein engineering steps required for target changes of ZNF and TALEN constructs. Moreover, the utilization of several sgRNAs in one reaction was shown to allow for multiplex editing of five genes [66].

All reported CRISPR genome editing approaches facilitate sgRNA:Cas9 systems to generate double strand breaks (DSB) of the target sequence, which can be either repaired by homologous recombination (HR) or non-homologous end joining (NHEJ) (Figure 2). In the case of HR, a full repair of the DSB in eukaryotes is facilitated, as the wild-type allele serves as a donor template. In contrast, NHEJ is an error-prone repair mechanism, which leads to the formation of insertions and deletions (indels) that can result in mutations of the particular sequence. The aim of genome editing using RGEN, ZNF or TALEN is to introduce such indels to mutate the gene of interest. However, HR is the preferred DSB repair mechanism *in vivo*, which leads to a lower efficiency of the editing process. To increase the overall yield of edited sequences, a modified donor DNA can be used, which serves as a template during HR (Figure 2).

The basic principles of established RGENs are similar for different cells. In all cases, the sequence of Cas9 was codon optimized for the particular expression system, and a eukaryotic nuclear localization signal (NLS) was fused to the protein. The plasmid encoded Cas9 production in human, mouse and yeast cells was realized by the *in vivo* production of the protein [58–66]. In zebrafish embryo cells, the mRNA of codon-optimized Cas9 fused to an NLS was microinjected before expression [60,62] (Figure 2). A variety of approaches were chosen to import the targeting sgRNA. One of the employed strategies was to transfect plasmids of a codon-optimized Cas9 together with RNase III, pre-crRNA and tracrRNA [58]. Here, it could be shown that host RNAses are sufficient for sgRNA production and that no additional RNase III enzyme is needed. In other experiments, plasmids coding for a single fusion sgRNA were successfully employed [61,64,65]. Additionally, the direct microinjection of a synthesized sgRNA into zebrafish cells or the transfection of *in vitro*-produced sgRNAs in human cells resulted in similar editing efficiencies [59,62]. For the bacterial systems of *E. coli* and *Streptococcus pneumoniae*, HR was exploited with donor template DNA. Here, genomic DNA containing a truncated type II CRISPR system combined with a modified donor DNA template was transformed to yield an edited genome [63,66]. The verification of a successful editing event in mammalian cells was based on the SURVEYOR assay in which a Cel-1 nuclease specifically cleaves mismatches in hybrids of wild-type and mutant PCR products of the targeted sequence region [58,59,62,64,65,69]. Alternatively, successful editing was followed by antibiotic selection in yeast and bacteria cells [61,63].

Figure 2. Targeted genome editing using a small guide RNA (sgRNA):Cas9 complex. A plasmid encoding the codon-optimized Cas9 (red) with a nuclear localization signal (NLS) and an sgRNA (yellow), including the desired targeting sequence, are transferred into the target cell. A functional sgRNA:Cas9 interference complex is assembled in the cell. A double strand breaks (DSB) at the targeted DNA sequence upstream of a PAM (green) is introduced by the sgRNA:Cas9 complex, which can be repaired by the host DNA repair mechanisms, homologous recombination (HR) and non-homologous end joining (NHEJ). While HR restores the wild-type sequence by using the template allele, the error-prone NHEJ mechanism leads to insertions and deletions (indels) at the target site (grey). Increased editing efficiency can be achieved by co-transferring a synthetic donor DNA template for a triggered HR (blue).
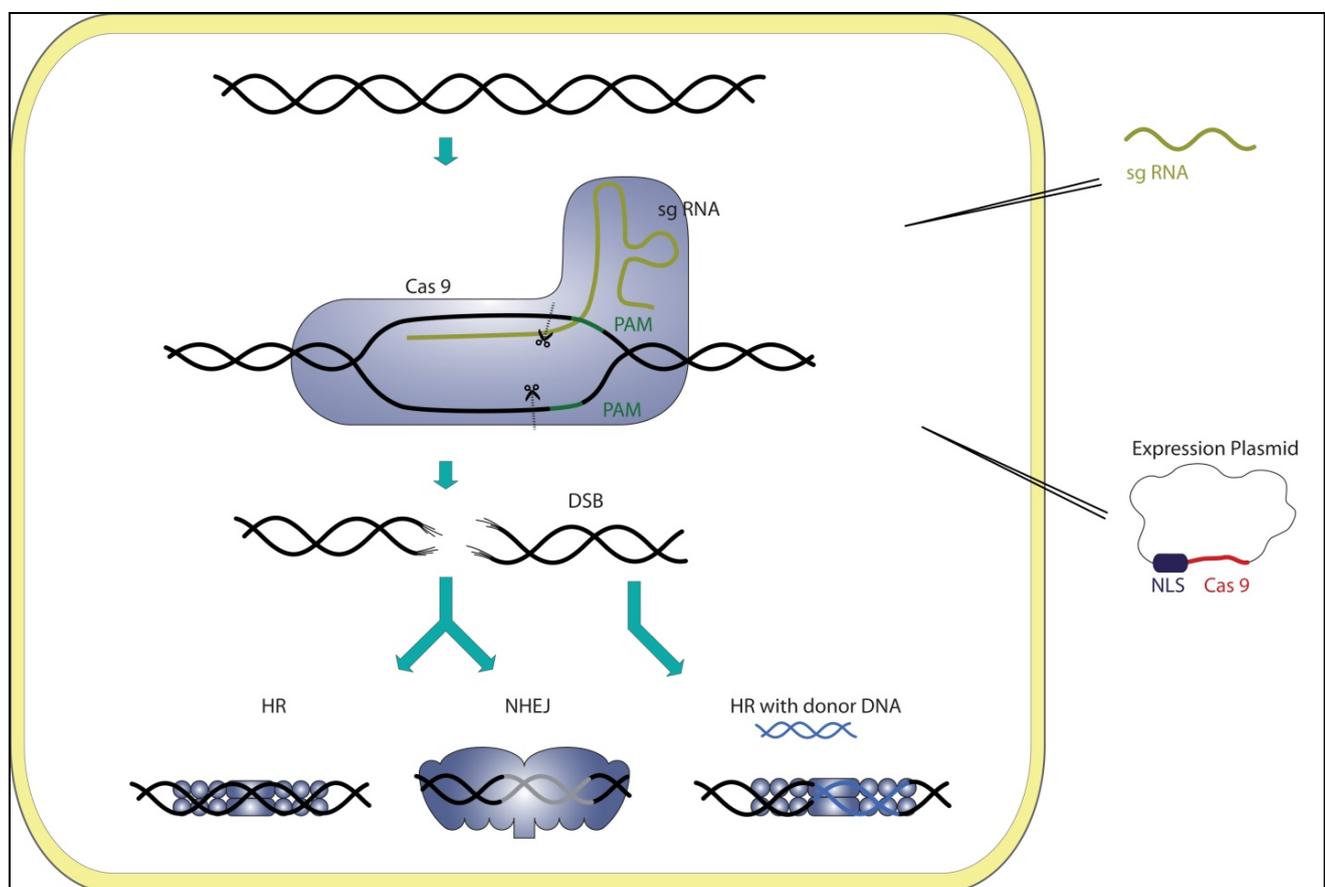
4. Type I and Type III Genome Editing Tools

Currently, only the type II CRISPR/Cas systems are established for genome editing or gene silencing [58–66]. What are the barriers that prevent the development of type I or type III editing tools? One difference between the three major CRISPR/Cas types is that the interference reaction of both, type I and III systems, relies on multi-protein complexes. This complicates the transfer of these systems to other organisms, and protein-engineering of a single Cas9 protein is more straightforward than optimization of Cascade or Cmr/Csm complexes would be [11,18]. However, the fact that type III systems do not need PAM sequences for interference should be advantageous for more versatile editing events [32]. One restriction of RGEN genome editing tools is the PAM sequence, which is

needed for interference. Therefore, this methodology enables the editing of sequences that occur on average every 8 bp, which leads to approximately 40% of the exons in the human genome being applicable for targeting [58,65,70]. Without the restricting PAM sequences for the interference in type III systems, genome editing could, in principle, be accomplished for any given target sequence. However, it should be noted that less sequence restriction needs to be in balance with specificity, as targeting of additional unwanted sequences has to be avoided. Similar to established bacterial editing systems, where truncated type II CRISPR/Cas systems were transformed, an approach using minimal type I (e.g., type I-F) or type III interference cassettes in combination with an engineered CRISPR cluster is conceivable. Nevertheless, the interference of a type I and, especially, type III CRISPR/Cas system and their minimal Cas protein assemblies are still not fully understood. Future research will address their potential role as genome editing tools.

5. Genome Editing Tools: RGEN, TALEN, ZFN

The advantages of RGENs over other genome editing tools rest upon the fast, simple and economical design of the targeting crRNAs and the usage of multiple small RNAs for multiplex engineering of several genes [58,62,63,65,66,70]. The ZiFiT Targeter program was designed as a tool to easily find potential sgRNA:Cas9 targeting sites in a given genome [62]. RGEN editing works at a similar efficiency compared to ZFN or TALEN, and so far, no off-site targeting was reported [59]. A further advantage of RGENs is that the CRISPR interference results in a selection for positive clones, as wild-type sequences are constantly targeted until the mutated donor DNA is integrated by HR and inhibits the interference process [61]. Other features of the RGEN approach reveal their restrictions and disadvantages in genome editing. One problem of RGENs is their dependence on a small PAM sequence, which restricts the choice of potential target sequences. Furthermore, the length of the crRNA and its seed sequence limit the range of target sequences and, thus, an optimal design of the sgRNA is required for efficient editing [59,62,64]. ZFN and TALEN systems require protein-engineering steps, but both systems are highly tunable by exchanging the particular domains for sequence specific recognition of the target. Furthermore, libraries of different proteins already exist [71], which reduces the time and costs for the generation of new functional proteins [59,72]. The modular structure of the two systems (TALEN and ZNF) yields a highly variable tool in which the effector domain can be fused to a variety of different proteins (e.g., transposases, nucleases, transcriptional regulators) [73–75]. ZFN- and TALEN-based gene therapy approaches are tested for clinical use. Further improvements, e.g., considering recent advances in re-organizing chromosomes using RGENs, will extend the number of possible applications in gene therapy [72,76].

6. Outlook

The potential of CRISPR/Cas systems to function in genetic tools has been discussed since the identification of the system as a prokaryotic immune system. Six years after its discovery, the first steps have been made, as sgRNA:Cas9 complexes are used for efficient genome editing [58–66] and the establishment of gene silencing [57]. Development of RGEN-based genome editing systems for further model organisms (e.g., plants, insects, Archaea) could simplify their future genetic manipulation. It is required to further increase the efficiency and elucidate the occurrence of off-site

targeting of RGENs to turn this system into a potential tool for disease treatment, as it has been, e.g., achieved with ZFN approaches for the treatment of HIV [72]. The establishment of RGEN genome editing tools is only one aspect, in which CRISPR/Cas immunity shows its potential for application in genetic and biotechnological systems. Other approaches include the development of phage-resistant bacterial strains used in industrial processes or the identification of the interaction of pathogenic bacteria in a clinical setting with either phages or conjugative plasmid that can transfer antibiotic resistance. Future research on basic mechanistic details of the different CRISPR/Cas systems will reveal a more complete picture of the extensive applicability of these immune complexes.

Acknowledgments

Conflict of Interest

The authors declare no conflict of interest.

References

1. Barrangou, R.; Fremaux, C.; Deveau, H.; Richards, M.; Boyaval, P.; Moineau, S.; Romero, D.A.; Horvath, P. CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **2007**, *315*, 1709–1712.
2. Grissa, I.; Vergnaud, G.; Pourcel, C. CRISPRFinder: A web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* **2007**, *35*, W52–W57.
3. Mojica, F.J.; Diez-Villasenor, C.; Garcia-Martinez, J.; Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **2005**, *60*, 174–182.
4. Pourcel, C.; Salvignol, G.; Vergnaud, G. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **2005**, *151*, 653–663.
5. Mojica, F.J.; Diez-Villasenor, C.; Garcia-Martinez, J.; Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **2009**, *155*, 733–740.
6. Swarts, D.C.; Mosterd, C.; van Passel, M.W.; Brouns, S.J. CRISPR interference directs strand specific spacer acquisition. *PLoS One* **2012**, *7*, e35888.
7. Pul, U.; Wurm, R.; Arslan, Z.; Geissen, R.; Hofmann, N.; Wagner, R. Identification and characterization of *E. coli* CRISPR-cas promoters and their silencing by H-NS. *Mol. Microbiol.* **2010**, *75*, 1495–1512.
8. Yosef, I.; Goren, M.G.; Qimron, U. Proteins and DNA elements essential for the CRISPR adaptation process in *Escherichia coli*. *Nucleic Acids Res.* **2012**, *40*, 5569–5576.

9. Babu, M.; Beloglazova, N.; Flick, R.; Graham, C.; Skarina, T.; Nocek, B.; Gagarinova, A.; Pogoutse, O.; Brown, G.; Binkowski, A.; *et al*. A dual function of the CRISPR-Cas system in bacterial antivirus immunity and DNA repair. *Mol. Microbiol.* **2011**, *79*, 484–502.

10. Diez-Villasenor, C.; Guzman, N.M.; Almendros, C.; Garcia-Martinez, J.; Mojica, F.J. CRISPR-spacer integration reporter plasmids reveal distinct genuine acquisition specificities among CRISPR-Cas I-E variants of *Escherichia coli. RNA Biol.* **2013**, *10*, doi:10.4161/rna.24023.

11. Hale, C.R.; Majumdar, S.; Elmore, J.; Pfister, N.; Compton, M.; Olson, S.; Resch, A.M.; Glover, C.V., III; Graveley, B.R.; Terns, R.M.; Terns, M.P. Essential features and rational design of CRISPR RNAs that function with the Cas RAMP module complex to cleave RNAs. *Mol. Cell* **2012**, *45*, 292–302.

12. Lillestol, R.K.; Shah, S.A.; Brugger, K.; Redder, P.; Phan, H.; Christiansen, J.; Garrett, R.A. CRISPR families of the crenarchaeal genus *Sulfolobus*: Bidirectional transcription and dynamic properties. *Mol. Microbiol.* **2009**, *72*, 259–272.

13. Carte, J.; Wang, R.; Li, H.; Terns, R.M.; Terns, M.P. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* **2008**, *22*, 3489–3496.

14. Gesner, E.M.; Schellenberg, M.J.; Garside, E.L.; George, M.M.; Macmillan, A.M. Recognition and maturation of effector RNAs in a CRISPR interference pathway. *Nat. Struct. Mol. Biol.* **2011**, *18*, 688–692.

15. Haurwitz, R.E.; Jinek, M.; Wiedenheft, B.; Zhou, K.; Doudna, J.A. Sequence- and structure-specific RNA processing by a CRISPR endonuclease. *Science* **2010**, *329*, 1355–1358.

16. Richter, H.; Zoephel, J.; Schermuly, J.; Maticzka, D.; Backofen, R.; Randau, L. Characterization of CRISPR RNA processing in *Clostridium thermocellum* and *Methanococcus maripaludis*. *Nucleic Acids Res.* **2012**, *40*, 9887–9896.

17. Sashital, D.G.; Jinek, M.; Doudna, J.A. An RNA-induced conformational change required for CRISPR RNA cleavage by the endoribonuclease Cse3. *Nat. Struct. Mol. Biol.* **2011**, *18*, 680–687.

18. 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*. Structural basis for CRISPR RNA-guided DNA recognition by Cascade. *Nat. Struct. Mol. Biol.* **2011**, *18*, 529–536.

19. 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.; van der Oost, J. Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **2008**, *321*, 960–964.

20. Wiedenheft, B.; Lander, G.C.; Zhou, K.; Jore, M.M.; Brouns, S.J.; van der Oost, J.; Doudna, J.A.; Nogales, E. Structures of the RNA-guided surveillance complex from a bacterial immune system. *Nature* **2011**, *477*, 486–489.

21. Wiedenheft, B.; van Duijn, E.; Bultema, J.B.; Waghmare, S.P.; Zhou, K.; Barendregt, A.; Westphal, W.; Heck, A.J.; Boekema, E.J.; Dickman, M.J.; Doudna, J.A. RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 10092–10097.

22. Bondy-Denomy, J.; Pawluk, A.; Maxwell, K.L.; Davidson, A.R. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **2013**, *493*, 429–432.

23. Haft, D.H.; Selengut, J.; Mongodin, E.F.; Nelson, K.E. A guild of 45 CRISPR-associated (Cas) protein families and multiple CRISPR/Cas subtypes exist in prokaryotic genomes. *PLoS Comput. Biol.* **2005**, *1*, e60.

24. Makarova, K.S.; Aravind, L.; Wolf, Y.I.; Koonin, E.V. Unification of Cas protein families and a simple scenario for the origin and evolution of CRISPR-Cas systems. *Biol. Direct* **2011**, *6*, doi:10.1186/1745-6150-6-38.

25. Makarova, K.S.; Haft, D.H.; Barrangou, R.; Brouns, S.J.; Charpentier, E.; Horvath, P.; Moineau, S.; Mojica, F.J.; Wolf, Y.I.; Yakunin, A.F.; *et al*. Evolution and classification of the CRISPR-Cas systems. *Nat. Rev. Microbiol.* **2011**, *9*, 467–477.

26. Beloglazova, N.; Petit, P.; Flick, R.; Brown, G.; Savchenko, A.; Yakunin, A.F. Structure and activity of the Cas3 HD nuclease MJ0384, an effector enzyme of the CRISPR interference. *EMBO J.* **2011**, *30*, 4616–4627.

27. Mulepati, S.; Bailey, S. Structural and biochemical analysis of nuclease domain of clustered regularly interspaced short palindromic repeat (CRISPR)-associated protein 3 (Cas3). *J. Biol. Chem.* **2011**, *286*, 31896–31903.

28. Sinkunas, T.; Gasiunas, G.; Fremaux, C.; Barrangou, R.; Horvath, P.; Siksnys, V. Cas3 is a single-stranded DNA nuclease and ATP-dependent helicase in the CRISPR/Cas immune system. *EMBO J.* **2011**, *30*, 1335–1342.

29. Westra, E.R.; Nilges, B.; van Erp, P.B.; van der Oost, J.; Dame, R.T.; Brouns, S.J. Cascade-mediated binding and bending of negatively supercoiled DNA. *RNA Biol.* **2012**, *9*, 1134–1138.

30. Sinkunas, T.; Gasiunas, G.; Waghmare, S.P.; Dickman, M.J.; Barrangou, R.; Horvath, P.; Siksnys, V. *In vitro* reconstitution of Cascade-mediated CRISPR immunity in *Streptococcus thermophilus*. *EMBO J.* **2013**, *32*, 385–394.

31. Semenova, E.; Jore, M.M.; Datsenko, K.A.; Semenova, A.; Westra, E.R.; Wanner, B.; van der Oost, J.; Brouns, S.J.; Severinov, K. Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 10098–10103.

32. Marraffini, L.A.; Sontheimer, E.J. Self *versus* non-self discrimination during CRISPR RNA-directed immunity. *Nature* **2010**, *463*, 568–571.

33. Sashital, D.G.; Wiedenheft, B.; Doudna, J.A. Mechanism of foreign DNA selection in a bacterial adaptive immune system. *Mol. Cell* **2012**, *46*, 606–615.

34. Nam, K.H.; Huang, Q.; Ke, A. Nucleic acid binding surface and dimer interface revealed by CRISPR-associated CasB protein structures. *FEBS Lett.* **2012**, *586*, 3956–3961.

35. Mulepati, S.; Orr, A.; Bailey, S. Crystal structure of the largest subunit of a bacterial RNA-guided immune complex and its role in DNA target binding. *J. Biol. Chem.* **2012**, *287*, 22445–22449.

36. Lintner, N.G.; Kerou, M.; Brumfield, S.K.; Graham, S.; Liu, H.; Naismith, J.H.; Sdano, M.; Peng, N.; She, Q.; Copie, V.; *et al*. Structural and functional characterization of an archaeal clustered regularly interspaced short palindromic repeat (CRISPR)-associated complex for antiviral defense (CASCADE). *J. Biol. Chem.* **2011**, *286*, 21643–21656.

37. Plagens, A.; Tjaden, B.; Hagemann, A.; Randau, L.; Hensel, R. Characterization of the CRISPR/Cas subtype I-A system of the hyperthermophilic crenarchaeon *Thermoproteus tenax*. *J. Bacteriol.* **2012**, *194*, 2491–2500.

38. Fischer, S.; Maier, L.K.; Stoll, B.; Brendel, J.; Fischer, E.; Pfeiffer, F.; Dyall-Smith, M.; Marchfelder, A. An archaeal immune system can detect multiple protospacer adjacent motifs (PAMs) to target invader DNA. *J. Biol. Chem.* **2012**, *287*, 33351–33363.

39. Peng, W.; Li, H.; Hallstrom, S.; Peng, N.; Liang, Y.X.; She, Q. Genetic determinants of PAM-dependent DNA targeting and pre-crRNA processing in *Sulfolobus islandicus*. *RNA Biol.* **2013**, *10*, doi:10.4161/rna.23798.

40. Sapranauskas, R.; Gasiunas, G.; Fremaux, C.; Barrangou, R.; Horvath, P.; Siksnys, V. The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* **2011**, *39*, 9275–9582.

41. Deltcheva, E.; Chylinski, K.; Sharma, C.M.; Gonzales, K.; Chao, Y.; Pirzada, Z.A.; Eckert, M.R.; Vogel, J.; Charpentier, E. CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. *Nature* **2011**, *471*, 602–607.

42. Chylinski, K.; Le Rhun, A.; Charpentier, E. The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems. *RNA Biol.* **2013**, *10*, doi:10.4161/rna.24321.

43. Jinek, M.; Chylinski, K.; Fonfara, I.; Hauer, M.; Doudna, J.A.; Charpentier, E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **2012**, *337*, 816–821.

44. Gasiunas, G.; Barrangou, R.; Horvath, P.; Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, E2579–E2586.

45. Garneau, J.E.; Dupuis, M.E.; Villion, M.; Romero, D.A.; Barrangou, R.; Boyaval, P.; Fremaux, C.; Horvath, P.; Magadan, A.H.; Moineau, S. The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **2010**, *468*, 67–71.

46. Marraffini, L.A.; Sontheimer, E.J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **2008**, *322*, 1843–1845.

47. Marraffini, L.A.; Sontheimer, E.J. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nat. Rev. Genetics* **2010**, *11*, 181–190.

48. Zhang, J.; Rouillon, C.; Kerou, M.; Reeks, J.; Brugger, K.; Graham, S.; Reimann, J.; Cannone, G.; Liu, H.; Albers, S.V.; *et al.* Structure and mechanism of the CMR complex for CRISPR-mediated antiviral immunity. *Mol. Cell* **2012**, *45*, 303–313.

49. Shao, Y.; Cocozaki, A.I.; Ramia, N.F.; Terns, R.M.; Terns, M.P.; Li, H. Structure of the Cmr2–Cmr3 subcomplex of the Cmr RNA silencing complex. *Structure* **2013**, *21*, 376–384.

50. Osawa, T.; Inanaga, H.; Numata, T. Crystal structure of the Cmr2-Cmr3 subcomplex in the CRISPR-Cas RNA silencing effector complex. *J. Mol. Biol.* **2013**, in press.

51. Hale, C.R.; Zhao, P.; Olson, S.; Duff, M.O.; Graveley, B.R.; Wells, L.; Terns, R.M.; Terns, M.P. RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **2009**, *139*, 945–956.

52. Deng, L.; Garrett, R.A.; Shah, S.A.; Peng, X.; She, Q. A novel interference mechanism by a type IIIB CRISPR-Cmr module in *Sulfolobus*. *Mol. Microbiol.* **2013**, *87*, 1088–1099.

53. Vergnaud, G.; Li, Y.; Gorge, O.; Cui, Y.; Song, Y.; Zhou, D.; Grissa, I.; Dentovskaya, S.V.; Platonov, M.E.; Rakin, A.; *et al*. Analysis of the three *Yersinia pestis* CRISPR loci provides new tools for phylogenetic studies and possibly for the investigation of ancient DNA. *Adv. Exp. Med. Biol.* **2007**, *603*, 327–338.

54. Dale, J.W.; Brittain, D.; Cataldi, A.A.; Cousins, D.; Crawford, J.T.; Driscoll, J.; Heersma, H.; Lillebaek, T.; Quitugua, T.; Rastogi, N.; *et al*. Spacer oligonucleotide typing of bacteria of the *Mycobacterium tuberculosis* complex: Recommendations for standardised nomenclature. *Int. J. Tuberc. Lung Dis.* **2001**, *5*, 216–219.

55. Qi, L.; Haurwitz, R.E.; Shao, W.; Doudna, J.A.; Arkin, A.P. RNA processing enables predictable programming of gene expression. *Nat. Biotechnol.* **2012**, *30*, 1002–1006.

56. Lee, H.Y.; Haurwitz, R.E.; Apffel, A.; Zhou, K.; Smart, B.; Wenger, C.D.; Laderman, S.; Bruhn, L.; Doudna, J.A. RNA-protein analysis using a conditional CRISPR nuclease. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 5416–5421.

57. Qi, L.S.; Larson, M.H.; Gilbert, L.A.; Doudna, J.A.; Weissman, J.S.; Arkin, A.P.; Lim, W.A. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **2013**, *152*, 1173–1183.

58. Cong, L.; Ran, F.A.; Cox, D.; Lin, S.; Barretto, R.; Habib, N.; Hsu, P.D.; Wu, X.; Jiang, W.; Marraffini, L.A.; Zhang, F. Multiplex genome engineering using CRISPR/Cas systems. *Science* **2013**, *339*, 819–823.

59. Cho, S.W.; Kim, S.; Kim, J.M.; Kim, J.S. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. *Nat. Biotechnol.* **2013**, *31*, 230–232.

60. Chang, N.; Sun, C.; Gao, L.; Zhu, D.; Xu, X.; Zhu, X.; Xiong, J.W.; Xi, J.J. Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. *Cell Res.* **2013**, *23*, 465–472.

61. Dicarlo, J.E.; Norville, J.E.; Mali, P.; Rios, X.; Aach, J.; Church, G.M. Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic Acids Res.* **2013**, *41*, 4336–4343.

62. Hwang, W.Y.; Fu, Y.; Reyon, D.; Maeder, M.L.; Tsai, S.Q.; Sander, J.D.; Peterson, R.T.; Yeh, J.R.; Joung, J.K. Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nat. Biotechnol.* **2013**, *31*, 227–229.

63. Jiang, W.; Bikard, D.; Cox, D.; Zhang, F.; Marraffini, L.A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* **2013**, *31*, 233–239.

64. Jinek, M.; East, A.; Cheng, A.; Lin, S.; Ma, E.; Doudna, J. RNA-programmed genome editing in human cells. *ELife* **2013**, *2*, e00471.

65. Mali, P.; Yang, L.; Esvelt, K.M.; Aach, J.; Guell, M.; DiCarlo, J.E.; Norville, J.E.; Church, G.M. RNA-guided human genome engineering via Cas9. *Science* **2013**, *339*, 823–826.

66. Wang, H.; Yang, H.; Shivalila, C.S.; Dawlaty, M.M.; Cheng, A.W.; Zhang, F.; Jaenisch, R. One-step generation of mice carrying mutations in multiple genes by CRISPR/cas-mediated genome engineering. *Cell* **2013**, *153*, 910–918.

67. Gratz, S.J.; Cummings, A.M.; Nguyen, J.N.; Hamm, D.C.; Donohue, L.K.; Harrison, M.M.; Wildonger, J.; O'Connor-Giles, K.M. Genome engineering of *Drosophila* with the CRISPR RNA-guided Cas9 nuclease. *Genetics* **2013**, doi:10.1534/genetics.113.152710.

68. Segal, D.J.; Meckler, J.F. Genome engineering at the dawn of the golden age. *Ann. Rev. Genomics Hum. Genet.* **2013**, doi:10.1146/annurev-genom-091212-153435.

69. Qiu, P.; Shandilya, H.; D'Alessio, J.M.; O'Connor, K.; Durocher, J.; Gerard, G.F. Mutation detection using Surveyor nuclease. *BioTechniques* **2004**, *36*, 702–707.

70. Horvath, P.; Barrangou, R. RNA-guided genome editing a la carte. *Cell Res.* **2013**, doi:10.1038/cr.2013.39.

71. Kim, Y.; Kweon, J.; Kim, A.; Chon, J.K.; Yoo, J.Y.; Kim, H.J.; Kim, S.; Lee, C.; Jeong, E.; Chung, E.; *et al*. A library of TAL effector nucleases spanning the human genome. *Nat. Biotechnol.* **2013**, *31*, 251–258.

72. Ramalingam, S.; Annaluru, N.; Chandrasegaran, S. A CRISPR way to engineer the human genome. *Genome Biol.* **2013**, *14*, doi:10.1186/gb-2013-14-2-107.

73. Mercer, A.C.; Gaj, T.; Fuller, R.P.; Barbas, C.F., III. Chimeric TALE recombinases with programmable DNA sequence specificity. *Nucleic Acids Res.* **2012**, *40*, 11163–11172.

74. Gaj, T.; Mercer, A.C.; Gersbach, C.A.; Gordley, R.M.; Barbas, C.F., III. Structure-guided reprogramming of serine recombinase DNA sequence specificity. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 498–503.

75. Yant, S.R.; Huang, Y.; Akache, B.; Kay, M.A. Site-directed transposon integration in human cells. *Nucleic Acids Res.* **2007**, *35*, e50.

76. Xiao, A.; Wang, Z.; Hu, Y.; Wu, Y.; Luo, Z.; Yang, Z.; Zu, Y.; Li, W.; Huang, P.; Tong, X.; *et al*. Chromosomal deletions and inversions mediated by TALENs and CRISPR/Cas in zebrafish. *Nucleic Acids Res.* **2013**, in press.