CRISPR/Cas9 as a Mutagenic Factor

Subjects: Genetics & Heredity

Contributor: Andrey R. Shumega , Youri I. Pavlov , Angelina V. Chirinskaite , Aleksandr A. Rubel , Sergey G. Inge-Vechtomov , Elena I. Stepchenkova

The discovery of the CRISPR/Cas9 microbial adaptive immune system has revolutionized the field of genetics, by greatly enhancing the capacity for genome editing. CRISPR/Cas9-based editing starts with DNA breaks (or other lesions) predominantly at target sites and, unfortunately, at off-target genome sites. DNA repair systems differing in accuracy participate in establishing desired genetic changes but also introduce unwanted mutations, that may lead to hereditary, oncological, and other diseases. New approaches to alleviate the risks associated with genome editing include attenuating the off-target activity of editing complex through the use of modified forms of Cas9 nuclease and single guide RNA (sgRNA), improving delivery methods for sgRNA/Cas9 complex, and directing DNA lesions caused by the sgRNA/Cas9 to non-mutagenic repair pathways.

genome editing CRISPR/Cas9 mutations off-target activity

1. Structure, Mechanism of Action and DNA Lesions Caused by CRISPR/Cas9 Editing Tools

Most CRISPR/Cas editing platforms originate from a natural system of adaptive immunity in bacteria and archaea responsible for cleaving foreign DNA that has entered the cell, e.g., plasmids or phages [1]. The formation of immunity, in this case, occurs in several steps. Upon entry, fragments of foreign DNA (known as spacers) are integrated into the CRISPR locus. When re-invasion occurs, transcription of the CRISPR locus begins, resulting in the production of a long molecule of pre-crRNA (pre-CRISPR-associated RNA), which is processed into short crRNAs (CRISPR-associated RNA). In the complex with Cas proteins, these short molecules bind to homologous segments of foreign DNA (known as protospacers), leading to the cleavage of the DNA by Cas proteins. Notably, the formation of a double-strand break in the protospacer DNA only occurs in the presence of a specific nucleotide sequence called the protospacer adjacent motif (PAM), which serves as a marker for foreign DNA [2].

All known natural CRISPR/Cas systems are divided into two classes and six types depending on the structure and mechanism of action [3]. The Class 1 systems are common in bacteria and are multi-protein types I, III, and IV complexes (Cascade, Cmr, Csm). The nuclease activity in Type I systems resides in the multi-protein Cascade complex associated with the Cas3 protein. Type III systems are characteristic of archaea, and nuclease activity is conferred by the multi-protein complexes Csm and Cmr. Type IV systems are relatively rare and poorly studied [4]. The Class 2 systems, on the other hand, have a single effector protein. This class includes types II, V, and VI. Type II systems are actively used in genetic engineering and are characterized by the presence of the endonuclease SpCas9 (*Streptococcus pyogenes*). SpCas9 (hereafter called Cas9) does not require additional protein cofactors

for binding and cleaving target DNA. Inside the cell, two RNA molecules are required to activate Cas9: the CRISPR-associated RNA (crRNA), which contains fragments of foreign sequences, and the trans-activating crRNA (tracrRNA), which base pairs with crRNA and supports crRNA maturation. Cr-tracrRNA complex activates Cas9 through conformational changes. However, for simplicity in genetic engineering applications, both RNAs are combined into a single chimeric molecule called a single-guide RNA (sgRNA) [2]. The nucleoprotein complex used for genome editing consists of a sgRNA and the nuclease Cas9 (**Figure 1**). The sgRNA contains a sequence of 19–20 nucleotides (guide sequence) that is complementary to the target region of the genome, as well as functional domains tracrRNA and crRNA (**Figure 1**A). It is important to note that a ten-nucleotide segment from the 3'-end of the guide sequence (seed region) plays a crucial role in the recognition and cleavage of the target DNA [5].

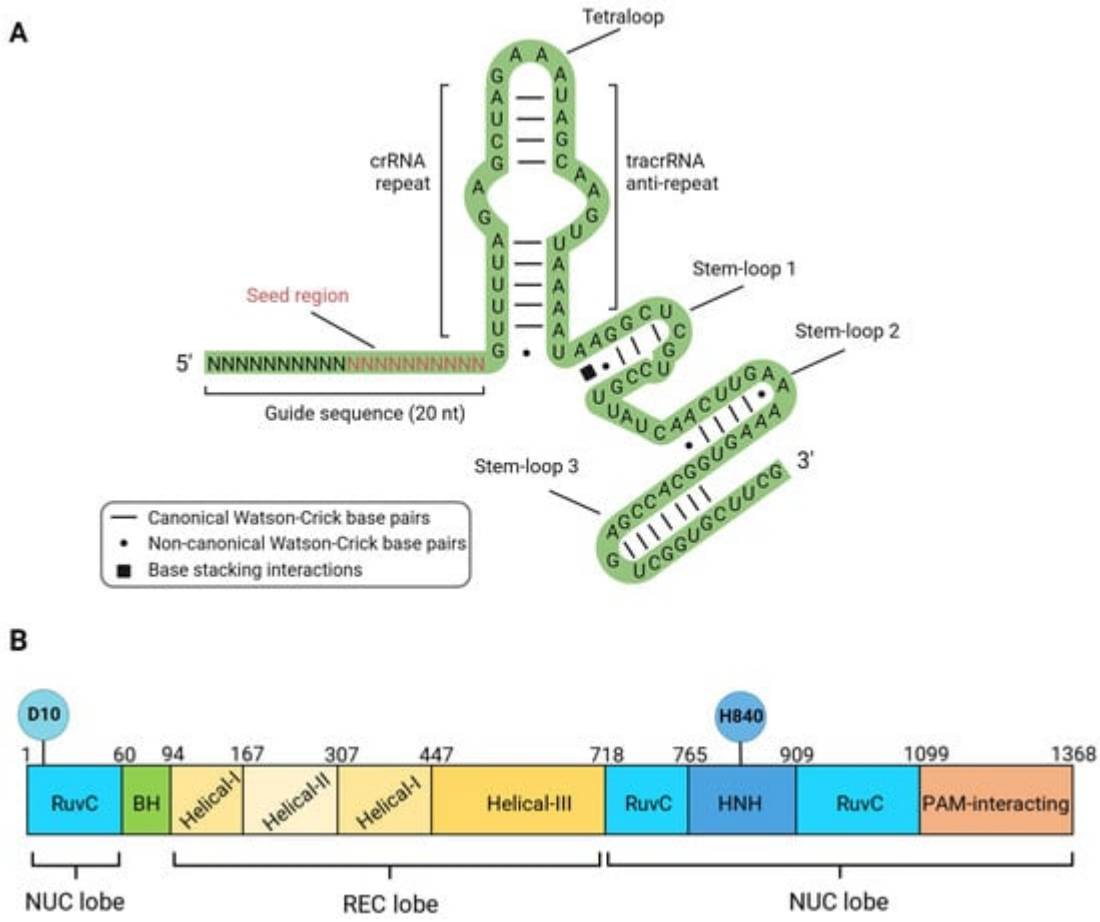

Figure 1. Components of active sgRNA/Cas9 editing complex. (**A**) Secondary structure of the sgRNA molecule: Guide sequence with 10 bp 3′ seed region; tracrRNA—trans-activating CRISPR RNA functional domain; crRNA—trans-activating CRISPR RNA functional domain. (**B**) Domain organization of the Cas9 protein. D10 and H840—amino acid residues essential for endonuclease catalytic activity. The nuclease (NUC) lobe is represented by endonuclease HNH and RuvC domains, the recognition (REC) lobe binds the nucleic acid and contains three recognition domains, Helical I–III.

The Cas9 protein consists of 10 domains, which are divided into two functional groups—the first group is responsible for DNA binding (helical recognition (REC) lobe), and the second group is responsible for introducing

double-strand breaks (nuclease (NUC) lobe). The REC lobe includes the domains Helical I, Helical II, and Helical III, while the NUC lobe includes the domains RuvC-I, RuvC-II, RuvC-III, HNH, and PAM-interacting domain CTD [2] (**Figure 1**B).

The search for the target DNA and introduction of double-strand breaks occur in several stages: (1) Assembly of the sgRNA/Cas9 complex, resulting in a conformational change in the central part of the molecule, opening up DNA binding domains; (2) Stochastic interactions between the sgRNA/Cas9 complex and chromosomal DNA, which must contain the PAM sequence of the NGG type directly adjacent to the 3′-end of the sgRNA guide sequence; (3) Formation of hydrogen bonds between the CTD domain of Cas9 and PAM sequences distributed throughout the genome, and unwinding of the DNA duplex one nucleotide away from the PAM; (4) Complementary sgRNA promotes further unwinding of the target DNA duplex; (5) RuvC and HNH domains of the Cas9 protein introduce a double-strand break with blunt ends at a distance of three nucleotides from the PAM sequence [2]. Thus, the CRISPR/Cas9 editing system possesses a property characteristic of most mutagens—the ability to damage DNA.

The nucleoprotein complex sgRNA/Cas9 is widely used as a platform for developing various genome editing tools that cause DNA lesions of different types. Along with the wild-type protein capable of inducing a double-strand break, mutant variants of the nuclease are broadly used (**Figure 1**B,E). Two endonuclease domains of Cas9 cleave different DNA strands at a distance of 3 nucleotides from the 5′-end of the PAM sequence. The Cas9 HNH nuclease domain cleaves the strand complementary to the sgRNA, while the RuvC-like domain cleaves the non-complementary strand. Independent cleavage of each DNA strand leads to the DSB appearance [2]. Substitutions D10A (RuvC domain) or H840A (HNH domain) lead to the loss of the corresponding endonuclease activity (**Figure 1**) [2]. Cas9 with one inactivated endonuclease domain is called Cas9 nickase (nCas9); it introduces single-strand breaks (SSBs), which could result in mutations in the target DNA (**Figure 1**B). Inactivation of both endonuclease domains results in a catalytically dead enzyme (dCas9), that does not cleave DNA but serves as a delivery platform for targeting other enzymes fused with dCas9.

Both dCas9 and nCas9 may be fused with different functional protein domains to induce directed base substitutions without DSB introduction. Such complex enzymes are called base editors. Base editors can be divided into four main groups, Cytosine Base Editors (CBE), Adenine Base Editors (ABE), Dual-Base Editors (Dual BE), and Prime Editors (PE), depending on the mechanisms of editing. CBEs mostly consist of dCas9 or nCas9 nuclease conjugated with cytosine deaminases of the AID/APOBEC family and uracil glycosylase inhibitor (UGI) [6][7][8][9][10][11][12]. Upon binding to the target DNA sequence, deamination of cytosines leads to uracil formation in DNA. CBEs deaminate C to U at a distance of ~15 bp from the PAM motif [6]. Additionally, the “activity window” of the base editor, typically 5 bp in length, may contain several C that can be deaminated to U, leading to multiple C to T substitutions [11]. nCas9 fused with TadA adenine deaminase with altered substrate specificity (RNA-specific enzyme turned to DNA-specific via directed evolution) is called ABE. Due to TadA activity, adenine is turned to inosine in the ~11 to 16 position from PAM. DNA base modifications in template DNA strands are highly mutagenic because they lead to base substitutions during DNA synthesis (**Figure 1**E). nCas9-base fused with TadA and APOBEC deaminases simultaneously is called Dual BE. Such editors have detectable off-target activity, similarly to

CBEs and ABEs, and editing efficiency is often lower than efficiency of CBE or ABE. Still, Dual BE is more versatile than CBE or ABE [13][14].

Each of Cas9 variants mentioned above is potentially mutagenic.

2. The Mechanism of Inaccurate Repair of Cas9-Induced DNA Lesions

There are two main mechanisms for repairing Cas9-induced DSBs—homology-dependent repair (HDR) and non-homologous end joining (NHEJ) (**Figure 1**A) [15][16][17]. However, minor pathways, such as microhomology-mediated end joining (MMEJ) [18] or break-induced replication (BIR) [19] may also be involved (**Figure 1**D). In general, the main factors determining the repair pathway are the DNA lesion's nature and the cell cycle's phase [20]. Repair of DSBs by NHEJ occurs during the G1 phase of the cell cycle, while HDR occurs during the G2/M phases. NHEJ often leads to genetic variability at Cas9 cleavage sites, as DNA exonucleases and low-fidelity DNA polymerases are involved in processing free DNA ends before ligation, increasing the likelihood of errors [21].

The initiation of NHEJ in mammalian cells occurs through binding to DNA ends in the double-strand break region by the Ku70/Ku80 heterodimer and DNA-dependent protein kinase (DNA-PK), which, together with the endonuclease Artemis, process the DNA ends (fragmentation, gap filling, removal of damaged nucleotides) for subsequent ligation. Then, the DNA ends are covalently joined by the DNA ligase IV/XRCC4 complex. Gap filling in DNA is carried out by polymerases λ and μ (PolX family), whose activation is mediated by the Ku/DNA complex through BRCT domains of the polymerases. Polymerases λ and μ are well suited for NHEJ as they are capable of synthesis without relying on a template [17].

In the case of homologous recombination, genetic information is restored using sister chromatids. In the presence of a transgenic DNA fragment with a locus homologous to a genomic region, recombination can occur between the transgene and the genomic DNA at the site of the double-strand break, resulting in the insertion of the transgenic sequence at the break site or replacement of the target sequence with the transgenic one. In the initial stage, end preparation is initiated at the double-strand break site by nucleotide excision in the 5'-3' direction with the formation of single-stranded DNA (ssDNA). Then, one of the strands of the homologous sequence invades, forming a Holliday junction structure. In yeast *Saccharomyces cerevisiae*, it has been shown that the MRX complex (homologous to the MRN complex in mammals), which includes proteins Mre11, Rad50, Xrs2, and Sae2, is involved in processing DSBs ends [22]. In mammalian cells, DNA-dependent ATPase protein Rad51 plays an important role in DNA strand exchange, forming nucleoprotein filaments with DNA [23] and catalyzing strand exchange by forming D-loops. It is known that the protein p53 is responsible for activating the repair of double-strand breaks in mammalian cells. Phosphorylation of p53 occurs in response to DNA damage and initiates the transcription of double-strand break repair genes (*BRCA2*, *RAD51*, and *MRE11*). Phosphorylation of p53 is carried out by checkpoint kinases ATM and ATR, with ATM being a sensor of double-strand breaks [24].

The effectiveness of introducing genetic changes using sgRNA/Cas9 largely depends on the balance between homologous and non-homologous DNA repair [25][26][27], which changes during the cell cycle. NHEJ factors are expressed in cells throughout the cell cycle, while the activation of factors involved in homologous recombination occurs in the S/G2 phases of the cell cycle by increased expression and post-translational modifications of specific proteins [28]. This makes the S/G2 phases of the cell cycle the most favorable time for introducing desired genetic changes, as there is a higher likelihood of homology-directed repair occurring. The low frequency of homologous recombination is a significant challenge in genetic modification. For example, due to the high sensitivity of human pluripotent stem cells to genome damage, the frequency of modifications using CRISPR/Cas9 systems is less than 10% [29]. Promising approaches to enhance the efficiency of site-directed mutagenesis using CRISPR/Cas9 systems include altering the balance between homologous and non-homologous repair and activating the homologous mechanism in atypical cell cycle phases. One way to suppress the NHEJ mechanism is to inhibit key proteins. When short hairpin RNAs (shRNAs) targeting Ku70, Ku80, and DNA ligase IV are used, the frequency of homologous recombination increases up to 300% [30]. Additionally, when a small subunit of the protein Scr7, which blocks the DNA-binding domain of DNA ligase IV, and proteins involved in proteasomal degradation of DNA ligase IV (E1B55K and E4orf6) are added to shRNAs, the frequency of site-directed mutagenesis increases seven-fold compared to the baseline level [31]. Inhibition of another component of NHEJ, DNA-dependent protein kinase catalytic subunit (DNA-PKcs), increases the frequency of site-directed mutagenesis four-fold [32].

The Rad51 protein is a crucial element of homologous recombination, forming a nucleofilament on single-stranded DNA and participating in DNA duplex invasion. Treatment of cells with the compound RS-1 (4-Bromo-*N*-(4-bromophenyl)-3-[[(phenylmethyl)amino]sulfonyl]benzamide) stabilizes the binding of Rad51 to DNA and increases the frequency of site-directed mutagenesis by three to six times [33]. The treatment of pluripotent stem cells with nocodazole or ABT-751 stimulates their transition into the G2/M phases of the cell cycle, which enhances homologous recombination and consequently increases the frequency of site-directed mutagenesis by four times [34].

Modifications of the CRISPR/Cas9 gene editing platform that prevent DSB formation and induce DNA lesions of different types (single-strand breaks (SSB) and base alterations) have fewer side effects and higher accuracy, but they are still genotoxic. nCas9 and PEs could be a source of all types of mutations due to HDR and single-strand break repair (SSBR) (**Figure 1**B) [35]. The base editors CBE, ABE, and Dual BE, causing deamination of cytosine and adenine, provoke base substitutions (C to T and A to G) due to replicative DNA synthesis on a template, containing deaminated bases (**Figure 1**E). Additionally, deamination of cytosines may occur spontaneously in R-loops, formed by sgRNA and genomic DNA during editing. Deaminated bases are substrates for BER and MMR (**Figure 1**C) [36], which are potentially mutagenic multistep processes. In the first step of BER, uracil or hypoxanthine is excised by specific DNA glycosylases generating apurinic/apyrimidinic (AP) sites. AP sites are then converted to SSBs with an undamaged 3′-OH terminus and a base-free 5′-deoxyribose 5-phosphate residue at the 5′ end. Such lesions may stimulate DSBs formation [37][38] and further processed either through short-patch or long-patch pathways mediated by different DNA polymerases. In the short-patch pathway, the main role in filling gaps belongs to Pol β, but Pol λ also contributes to BER. Long-patch BER is mediated by Pol δ, or ε, that have high fidelity and fill longer gaps [39]. During MMR, single stand gaps are formed after the excision of unpaired or

damaged nucleotides. Then, they are filled by Pol δ, or ε. Intermediates of SSBR, BER, and MMR are mutagenic because of limited fidelity of DNA polymerases, participating in gap filling during repair. Moreover, such BER intermediates, as AP sites are bypassed by TransLesion Synthesis (TLS) DNA polymerases, during replication when replication forks are stalled at the lesion. TLS is an error-prone process that leads to an increase in the rates of base substitutions, short insertions and deletions, complex mutations, and long deletions between repeating sequences [40][41][42]. By using different forms of Cas9 that cause specific types of DNA lesions, it is possible to shift the repair process toward a specific repair pathway and increase the likelihood of obtaining certain types of mutations. For example, it is well known that short insertions and deletions are the most common result of repair of Cas9-induced double-strand breaks. For the induction of longer deletions, Cas9 fused with human APOBEC3A, uracil DNA glucosidase, and apurinic or apyrimidinic site lyase may be used [43]. All the scenarios may take place at on- and off-target sites.

3. On-Target and Off-Target Activity: Hot Points of the sgRNA/Cas9-Dependent Mutagenesis

One of the important features of the sgRNA/Cas9 mutagenic activity is its high sequence specificity compared to other mutagenic factors. The high site-specific activity is determined by the presence of a PAM sequence and the subsequent 20 nucleotides complementary to the artificial sgRNA. The complex efficiency depends on the nucleotide at position N for the SpCas9 PAM (NGG), as it affects the change in the complex's binding free energy with DNA and the stability of the sgRNA/Cas9 complex on genomic DNA. Therefore, a particular nucleotide at position N can have different effects on the activity of the editor [44]. However, it should be noted that the specificity of sgRNA/Cas9 is not absolute. The non-specific activity of sgRNA/Cas9 in off-target regions of the genome is one of the urgent problems, as it can lead to the emergence of pathogenic mutations, for example, in oncogenes or tumor suppressor genes. It is unclear why some off-target sites are cleaved by the Cas9 protein while others are not. The efficiency of sgRNA/Cas9 action can be affected by the chromatin structure of the target locus and by the presence of epigenetic modifications, such as DNA methylation or histone modifications [45][46][47]. In addition, the cell cycle stage and the level of DNA damage response can also affect the efficiency and specificity of sgRNA/Cas9 cleavage [48].

The level of expression and duration of action of editors in cells strongly influence the activity and accuracy of gene correction. Therefore, the delivery methods of sgRNA/Cas9 to target cells greatly influence its off-target effect [49][50][51][52][53][54]. Cas9 can be delivered into cells in the forms of DNA (plasmid transfection, viral transduction) or mRNA (microinjections, electroporation, or liposome-based vectors [55][56][57][58]), or it can be delivered into cells as sgRNA/protein complex (electroporation). In the latter case, the off-target activity is lower than in other delivery methods. For example, plasmid DNA poses an additional risk of insertional mutagenesis [59]. The expression of AAV-delivered genes (adeno-associated virus) persists for years in transfected cells [60][61]. Thus, AAV-mediated gene editing will likely yield unwanted on- and off-target mutations over time [62][63]. In contrast, Cas9 mRNA and sgRNA delivered by lipid nanoparticles (LNP) can be rapidly degraded in vivo, making LNP the most popular vector for in vivo gene editing [64].

Thus, for the sgRNA/Cas9 editing complex, hotspots of mutagenesis are sequences that are fully or partially complementary to the sgRNA. It is important to note that different sgRNAs, which are fully complementary to different sites in the genomic DNA differ in their ability to bind to the target sequence and, as a result, in the frequency of induced mutations. Since the delivery method and duration of exposure of the editing complex affect the frequency of mutagenesis, it can be concluded that this mutagenic factor, like other mutagens, exhibits a dose-dependent relationship.

The choice of guide sequence and the sgRNA/Cas9 complex's design can influence its activity and specificity. The use of modified single guide RNAs or optimizing the PAM sequence can increase the specificity of sgRNA/Cas9 editing, while using modified Cas9 enzymes with reduced off-target activity can further improve its safety and accuracy. Various modifications of the Cas9 protein are often used to increase the activity and accuracy of editing. By using a nCas9 with two correctly matched sgRNAs, it is possible to obtain two close single-strand breaks that together produce a double-strand break [65][66]. The increase in specificity in the case of using paired sgRNA/Cas9-D10A complexes is about two orders of magnitude in the HEK-293FT cell line [66]. Increased specificity is also observed for protein variants eSpCas(1.0) (K810A K1003A R1060A) and eSpCas(1.1) (K848A K1003A R1060A). They made no off-target changes at 22 of the 24 predicted most likely off-target sites and were sensitive to one sgRNA mismatch outside the seed sequence [67]. In the highly specific nuclease SpCas9-HF1 with substitutions N497A R661A Q695A Q926A, the frequency of off-target mutations induced by SpCas9-HF1 was not statistically different from the spontaneous level of mutagenesis [68]. HypaCas9 variant (N692A M694A Q695A H698A) with mutations in the REC3 domain, is more specific compared to wt Cas9 and even to eSpCas(1.1) and SpCas9-HF1 and preserves similar efficiency [69]. Another variant, SuperFi-Cas9 (Y1010D Y1013D Y1016D V1018D R1019D Q1027D K1031D) showed a 500-fold increase in specificity [70], but its effect in cells has not yet been studied. Increased accuracy is also observed in enzymes evoCas9 (M495V Y515N K526E R661Q) [71], Sniper-Cas9 (F539S M763I K890N) [72] and HiFi Cas9 (R691A) [73]. Variants with multiple amino acid substitutions xCas9-3.6 E108G S217A A262T S409I E480K E543D M694I E1219V and xCas9-3.7 A262T R324L S409I E480K E543D M694 I E1219V, recognizing such PAM motifs as NG, NNG, GAA, GAT, and CAA, additionally exhibited in 10–100-fold higher target specificity in HEK-293T and U2OS cell lines [74].

Another factor influencing the non-specific activity of sgRNA/Cas9 is the number of nucleotide mismatches between the sgRNA guide sequence and the target chromosomal DNA, as well as the positions of the mismatches relative to the PAM sequence. It is assumed that the number of non-complementary bases between the guide sequence and a potential off-target sequence significantly affects off-target activity. The sgRNA/Cas9 complex can introduce double-strand breaks into sequences with incomplete homology to the guide sequence, containing up to four mismatched nucleotides [75] or even up to six contiguous mismatches [2]. Mismatches within a short sequence (8 to 12 nucleotides proximal to PAM) rarely influence off-target activity of the editor, while more distal mismatches relative to PAM lead to an increase in the frequency of off-target activity [76]. In other words, off-target activity correlates with the stability of the sgRNA/Cas9 complex on genomic DNA. An excess of potential energy in the interaction between the Cas9 protein and the PAM sequence of the target site can stabilize the Cas9-sgRNA-DNA complex when binding to an off-target site containing non-complementary bases. Firstly, it has been shown that the nuclease activity of Cas9 is activated after DNA strand unwinding at the target site [77]. In addition, the presence of

non-complementary bases between the sgRNA guide sequence and the target DNA site within 1–12 nucleotides proximal to PAM suppresses the nuclease activity of Cas9. Secondly, certain nucleotides at positions 2, 3, 6, from the 5′ end of the sgRNA also negatively affect the activity of the complex. The 20th nucleotide from the 5′ end of the sgRNA preceding the PAM sequence is involved in initiating the unwinding of the DNA strand and stabilizing the Cas9-sgRNA-DNA complex in the PAM region. sgRNA molecules containing adenine at this position show a significant decrease in activity. Some nucleotides at positions 2 (Thymine), 3 (Guanine), and 6 (Adenine) also reduce the activity of the complex, despite the assumption that the 5′ end of the sgRNA (distal to PAM) is not essential for recognizing the genomic target site. Presumably, the decrease in complex binding efficiency to DNA is associated with weakened interaction between nucleotides 2, 3, and 6 and the Rec1 domain of the Cas9 protein [29]. It is also known that the efficiency of double-strand breaks is positively influenced by the optimal GC composition of the target sequence (40–60%), the ability to form secondary structures, and chromatin activity near the break (promoter regions) [29].

It has been shown that altering the secondary structure of the sgRNA molecule by adding two additional guanine nucleotides at the 5′ end leads to a decrease in non-specific activity without reducing target activity [78]. The addition of several cytosine nucleotides to the 5’ end of sgRNA significantly modifies the activity of the nucleoprotein complex. Polycytosine tails, depending on their length, reduce the cytotoxicity of the editing complex due to the low activation of p53, leading to the stimulation of homologous repair or a reduction in the on-target activity of the nuclease, improving the specificity and accuracy of monoallelic editing [79]. Shortening the sgRNA by three nucleotides leads to a decrease in the potential energy of complex binding to DNA, which significantly reduces off-target activity [80]. Chemical modifications of CRISPR RNA (crRNA) (**Figure 1**A) are also promising. As with other nucleic acid-based technologies, efforts are focused on sugar and backbone modifications (2′-deoxy, 2′-F, 2′-OMe, and phosphorothioates). Some more significant modifications of crRNAs have been made using bicyclic (locked) ribose and phosphate backbone substitutions (phosphonoacetates and amides); however, the range of chemical modifications applied to crRNA remains limited to modifications that have been successful in RNA interference and antisense technologies. Encouraging results on editing efficiency and accuracy have been obtained [81]. All of the above observations are actively used to design gene editing systems aimed at increasing efficiency and reducing off-target activity, including new target and guide sequence selection algorithms, new transfection methods, and new variants of Cas9/sgRNA modifications (**Figure 2**).

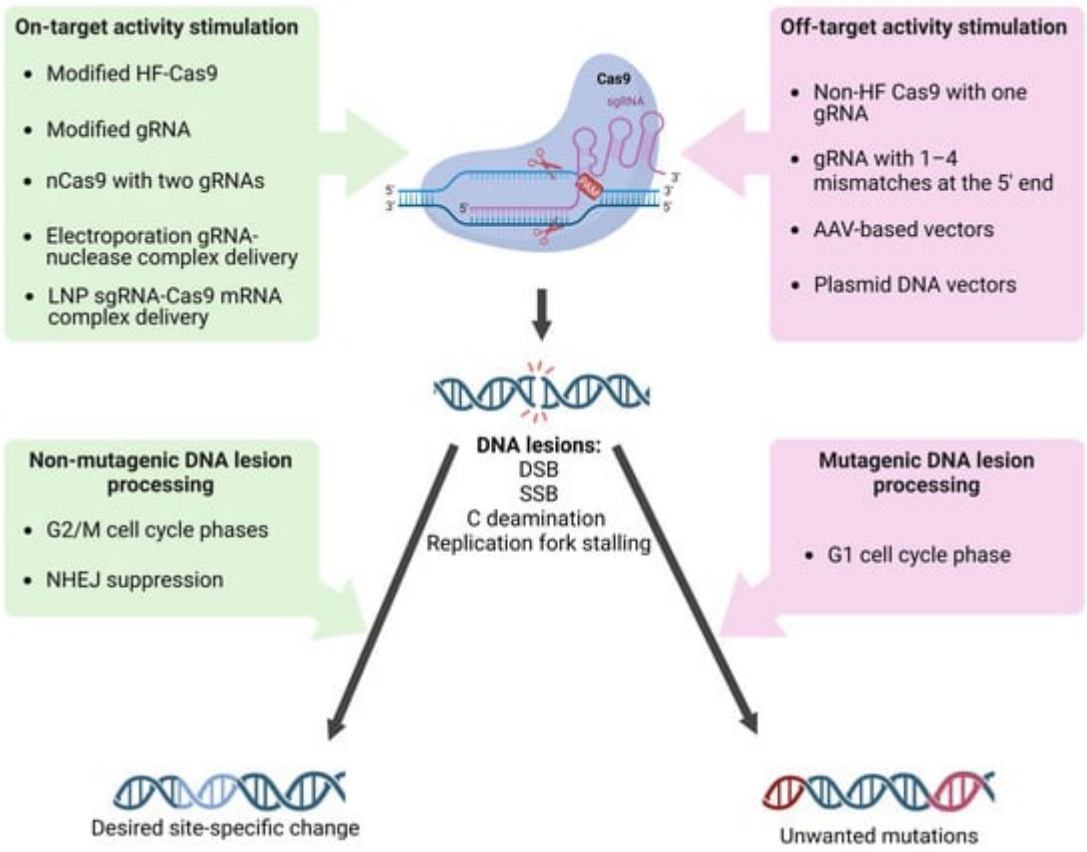

Figure 2. Factors influencing sgRNA/Cas9 targeting and DNA damage processing fidelity. The most accurate editing occurs when the factors listed on the left operate sequentially. Combinations of factors (on the right) that reduce the specificity of sgRNA/Cas9 binding to DNA and factors that direct lesions to mutagenic processing lead to the accumulation of unwanted mutations both in off- and target sites.

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. 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.
3. Makarova, K.S.; Wolf, Y.I.; Iranzo, J.; Shmakov, S.A.; Alkhnbashi, O.S.; Brouns, S.J.J.; Charpentier, E.; Cheng, D.; Haft, D.H.; Horvath, P.; et al. Evolutionary Classification of CRISPR-Cas Systems: A Burst of Class 2 and Derived Variants. Nat. Rev. Microbiol. 2020, 18, 67–83.
4. Barrangou, R. Diversity of CRISPR-Cas Immune Systems and Molecular Machines. Genome Biol. 2015, 16, 247.

5. Wong, N.; Liu, W.; Wang, X. WU-CRISPR: Characteristics of Functional Guide RNAs for the CRISPR/Cas9 System. Genome Biol. 2015, 16, 218.
6. Nishida, K.; Arazoe, T.; Yachie, N.; Banno, S.; Kakimoto, M.; Tabata, M.; Mochizuki, M.; Miyabe, A.; Araki, M.; Hara, K.Y.; et al. Targeted Nucleotide Editing Using Hybrid Prokaryotic and Vertebrate Adaptive Immune Systems. Science 2016, 353, aaf8729.
7. Skrekas, C.; Limeta, A.; Siewers, V.; David, F. Targeted In Vivo Mutagenesis in Yeast Using CRISPR/Cas9 and Hyperactive Cytidine and Adenine Deaminases. ACS Synth. Biol. 2023, 12, 2278–2289.
8. Chen, L.; Zhu, B.; Ru, G.; Meng, H.; Yan, Y.; Hong, M.; Zhang, D.; Luan, C.; Zhang, S.; Wu, H.; et al. Re-Engineering the Adenine Deaminase TadA-8e for Efficient and Specific CRISPR-Based Cytosine Base Editing. Nat. Biotechnol. 2023, 41, 663–672.
9. Zhang, Y.; Zhang, H.; Wang, Z.; Wu, Z.; Wang, Y.; Tang, N.; Xu, X.; Zhao, S.; Chen, W.; Ji, Q. Programmable Adenine Deamination in Bacteria Using a Cas9-Adenine-Deaminase fusion. Chem. Sci. 2020, 11, 1657–1664.
10. Wei, Y.; Zhang, X.H.; Li, D.L. The “New Favorite” of Gene Editing Technology-Single Base Editors. Yi Chuan 2017, 39, 1115–1121.
11. Komor, A.C.; Kim, Y.B.; Packer, M.S.; Zuris, J.A.; Liu, D.R. Programmable Editing of a Target Base in Genomic DNA without Double-Stranded DNA Cleavage. Nature 2016, 533, 420–424.
12. Komor, A.C.; Zhao, K.T.; Packer, M.S.; Gaudelli, N.M.; Waterbury, A.L.; Koblan, L.W.; Kim, Y.B.; Badran, A.H.; Liu, D.R. Improved Base Excision Repair Inhibition and Bacteriophage Mu Gam Protein Yields C:G-to-T:A Base Editors with Higher Efficiency and Product Purity. Sci. Adv. 2017, 3, eaao4774.
13. Grunewald, J.; Zhou, R.; Lareau, C.A.; Garcia, S.P.; Iyer, S.; Miller, B.R.; Langner, L.M.; Hsu, J.Y.; Aryee, M.J.; Joung, J.K. A Dual-Deaminase CRISPR Base Editor Enables Concurrent Adenine and Cytosine Editing. Nat. Biotechnol. 2020, 38, 861–864.
14. Zhao, Y.; Li, M.; Liu, J.; Xue, X.; Zhong, J.; Lin, J.; Ye, B.; Chen, J.; Qiao, Y. Dual Guide RNA-Mediated Concurrent C&G-to-T&A and A&T-to-G&C Conversions Using CRISPR Base Editors. Comput. Struct. Biotechnol. J. 2023, 21, 856–868.
15. Heyer, W.D.; Ehmsen, K.T.; Liu, J. Regulation of Homologous Recombination in Eukaryotes. Annu. Rev. Genet. 2010, 44, 113–139.
16. Marini, F.; Rawal, C.C.; Liberi, G.; Pellicioli, A. Regulation of DNA Double Strand Breaks Processing: Focus on Barriers. Front. Mol. Biosci. 2019, 6, 55.
17. Lieber, M.R. The mechanism of double-strand DNA Break Repair by the Nonhomologous DNA end-joining pathway. Annu. Rev. Biochem. 2010, 79, 181–211.

18. Sfeir, A.; Symington, L.S. Microhomology-Mediated End Joining: A Back-up Survival Mechanism or Dedicated Pathway? Trends Biochem. Sci. 2015, 40, 701–714.
19. Liu, L.; Malkova, A. Break-Induced Replication: Unraveling Each Step. Trends Genet. 2022, 38, 752–765.
20. Zhuk, A.S.; Shiriaeva, A.A.; Andreychuk, Y.V.; Kochenova, O.V.; Tarakhovskaya, E.R.; Bure, V.M.; Pavlov, Y.I.; Inge-Vechtomov, S.G.; Stepchenkova, E.I. Detection of Primary DNA Lesions by Transient Changes in Mating Behavior in Yeast Saccharomyces cerevisiae Using the Alpha-Test. Int. J. Mol. Sci. 2023, 24, 12163.
21. Geisinger, J.M.; Turan, S.; Hernandez, S.; Spector, L.P.; Calos, M.P. In Vivo Blunt-End Cloning through CRISPR/Cas9-Facilitated Non-Homologous End-Joining. Nucleic Acids Res. 2016, 44, e76.
22. Chapman, J.R.; Taylor, M.R.; Boulton, S.J. Playing the End Game: DNA Double-Strand Break Repair Pathway Choice. Mol. Cell 2012, 47, 497–510.
23. Li, X.; Heyer, W.D. Homologous Recombination in DNA Repair and DNA Damage Tolerance. Cell Res. 2008, 18, 99–113.
24. Abuetabh, Y.; Wu, H.H.; Chai, C.; Al Yousef, H.; Persad, S.; Sergi, C.M.; Leng, R. DNA Damage Response Revisited: The p53 Family and Its Regulators Provide Endless Cancer Therapy Opportunities. Exp. Mol. Med. 2022, 54, 1658–1669.
25. Hussmann, J.A.; Ling, J.; Ravisankar, P.; Yan, J.; Cirincione, A.; Xu, A.; Simpson, D.; Yang, D.; Bothmer, A.; Cotta-Ramusino, C.; et al. Mapping the Genetic Landscape of DNA Double-Strand Break Repair. Cell 2021, 184, 5653–5669.e25.
26. Xue, C.; Greene, E.C. DNA Repair Pathway Choices in CRISPR-Cas9-Mediated Genome Editing. Trends Genet. 2021, 37, 639–656.
27. Alkan, F.; Wenzel, A.; Anthon, C.; Havgaard, J.H.; Gorodkin, J. CRISPR-Cas9 off-Targeting Assessment with Nucleic Acid Duplex Energy Parameters. Genome Biol. 2018, 19, 177.
28. Mathiasen, D.P.; Lisby, M. Cell Cycle Regulation of Homologous Recombination in Saccharomyces cerevisiae. FEMS Microbiol. Rev. 2014, 38, 172–184.
29. Liu, X.; Homma, A.; Sayadi, J.; Yang, S.; Ohashi, J.; Takumi, T. Sequence Features Associated with the Cleavage Efficiency of CRISPR/Cas9 System. Sci. Rep. 2016, 6, 19675.
30. Chu, V.T.; Weber, T.; Wefers, B.; Wurst, W.; Sander, S.; Rajewsky, K.; Kuhn, R. Increasing the Efficiency of Homology-Directed Repair for CRISPR-Cas9-Induced Precise Gene Editing in Mammalian Cells. Nat. Biotechnol. 2015, 33, 543–548.
31. Maruyama, T.; Dougan, S.K.; Truttmann, M.C.; Bilate, A.M.; Ingram, J.R.; Ploegh, H.L. Increasing the Efficiency of Precise Genome Editing with CRISPR-Cas9 by Inhibition of Nonhomologous End

Joining. Nat. Biotechnol. 2015, 33, 538–542.

32. Robert, F.; Barbeau, M.; Ethier, S.; Dostie, J.; Pelletier, J. Pharmacological Inhibition of DNA-PK Stimulates Cas9-Mediated Genome Editing. Genome Med. 2015, 7, 93.
33. Pinder, J.; Salsman, J.; Dellaire, G. Nuclear Domain ‘Knock-in’ Screen for the Evaluation and Identification of Small Molecule Enhancers of CRISPR-Based Genome Editing. Nucleic Acids Res. 2015, 43, 9379–9392.
34. Yang, D.; Scavuzzo, M.A.; Chmielowiec, J.; Sharp, R.; Bajic, A.; Borowiak, M. Enrichment of G2/M Cell Cycle Phase in Human Pluripotent Stem Cells Enhances HDR-Mediated Gene Repair with Customizable Endonucleases. Sci. Rep. 2016, 6, 21264.
35. Caldecott, K.W. Single-Strand Break Repair and Genetic Disease. Nat. Rev. Genet. 2008, 9, 619–631.
36. Park, S.; Beal, P.A. Off-Target Editing by CRISPR-Guided DNA Base Editors. Biochemistry 2019, 58, 3727–3734.
37. Kouzminova, E.A.; Kuzminov, A. Fragmentation of Replicating Chromosomes Triggered by Uracil in DNA. J. Mol. Biol. 2006, 355, 20–33.
38. Kuzminov, A. Single-Strand Interruptions in Replicating Chromosomes Cause Double-Strand Breaks. Proc. Natl. Acad. Sci. USA 2001, 98, 8241–8246.
39. Krokan, H.E.; Bjoras, M. Base Excision Repair. Cold Spring Harb. Perspect. Biol. 2013, 5, a012583.
40. Northam, M.R.; Robinson, H.A.; Kochenova, O.V.; Shcherbakova, P.V. Participation of DNA Polymerase Zeta in Replication of Undamaged DNA in Saccharomyces cerevisiae. Genetics 2010, 184, 27–42.
41. Northam, M.R.; Moore, E.A.; Mertz, T.M.; Binz, S.K.; Stith, C.M.; Stepchenkova, E.I.; Wendt, K.L.; Burgers, P.M.; Shcherbakova, P.V. DNA Polymerases Zeta and Rev1 Mediate Error-Prone bypass of Non-B DNA Structures. Nucleic Acids Res. 2014, 42, 290–306.
42. Stepchenkova, E.I.; Zadorsky, S.P.; Shumega, A.R.; Aksenova, A.Y. Practical Approaches for the Yeast Saccharomyces cerevisiae Genome Modification. Int. J. Mol. Sci. 2023, 24, 11960.
43. Wang, S.; Zong, Y.; Lin, Q.; Zhang, H.; Chai, Z.; Zhang, D.; Chen, K.; Qiu, J.L.; Gao, C. Precise, Predictable Multi-Nucleotide Deletions in Rice and Wheat Using APOBEC-Cas9. Nat. Biotechnol. 2020, 38, 1460–1465.
44. Corsi, G.I.; Qu, K.; Alkan, F.; Pan, X.; Luo, Y.; Gorodkin, J. CRISPR/Cas9 gRNA Activity Depends on Free Energy Changes and on the Target PAM Context. Nat. Commun. 2022, 13, 3006.

45. Daer, R.M.; Cutts, J.P.; Brafman, D.A.; Haynes, K.A. The Impact of Chromatin Dynamics on Cas9-Mediated Genome Editing in Human Cells. ACS Synth. Biol. 2017, 6, 428–438.
46. Chen, E.; Lin-Shiao, E.; Trinidad, M.; Saffari Doost, M.; Colognori, D.; Doudna, J.A. Decorating Chromatin for Enhanced Genome Editing Using CRISPR-Cas9. Proc. Natl. Acad. Sci. USA 2022, 119, e2204259119.
47. Schep, R.; Brinkman, E.K.; Leemans, C.; Vergara, X.; van der Weide, R.H.; Morris, B.; van Schaik, T.; Manzo, S.G.; Peric-Hupkes, D.; van den Berg, J.; et al. Impact of Chromatin Context on Cas9-Induced DNA Double-Strand Break Repair Pathway Balance. Mol. Cell 2021, 81, 2216–2230.e10.
48. Nambiar, T.S.; Baudrier, L.; Billon, P.; Ciccia, A. CRISPR-Based Genome Editing through the Lens of DNA Repair. Mol. Cell 2022, 82, 348–388.
49. Lohia, A.; Sahel, D.K.; Salman, M.; Singh, V.; Mariappan, I.; Mittal, A.; Chitkara, D. Delivery Strategies for CRISPR/Cas Genome Editing Tool for Retinal Dystrophies: Challenges and Opportunities. Asian J. Pharm. Sci. 2022, 17, 153–176.
50. Carpenter, M.A.; Law, E.K.; Serebrenik, A.; Brown, W.L.; Harris, R.S. A Lentivirus-Based System for Cas9/gRNA Expression and Subsequent Removal by Cre-Mediated Recombination. Methods 2019, 156, 79–84.
51. Kim, S.; Kim, D.; Cho, S.W.; Kim, J.; Kim, J.S. Highly Efficient RNA-Guided Genome Editing in Human Cells via Delivery of Purified Cas9 Ribonucleoproteins. Genome Res. 2014, 24, 1012–1019.
52. Ramakrishna, S.; Kwaku Dad, A.B.; Beloor, J.; Gopalappa, R.; Lee, S.K.; Kim, H. Gene Disruption by Cell-Penetrating Peptide-Mediated Delivery of Cas9 Protein and Guide RNA. Genome Res. 2014, 24, 1020–1027.
53. Qiu, M.; Glass, Z.; Chen, J.; Haas, M.; Jin, X.; Zhao, X.; Rui, X.; Ye, Z.; Li, Y.; Zhang, F.; et al. Lipid Nanoparticle-Mediated Codelivery of Cas9 mRNA and Single-Guide RNA Achieves Liver-Specific In Vivo Genome Editing of Angptl3. Proc. Natl. Acad. Sci. USA 2021, 118, e2020401118.
54. Taha, E.A.; Lee, J.; Hotta, A. Delivery of CRISPR-Cas Tools for In Vivo Genome Editing Therapy: Trends and Challenges. J. Control. Release 2022, 342, 345–361.
55. Sinclair, F.; Begum, A.A.; Dai, C.C.; Toth, I.; Moyle, P.M. Recent Advances in the Delivery and Applications of Nonviral CRISPR/Cas9 Gene Editing. Drug Deliv. Transl. Res. 2023, 13, 1500–1519.
56. Yu, X.; Yang, Z.; Zhang, Y.; Xia, J.; Zhang, J.; Han, Q.; Yu, H.; Wu, C.; Xu, Y.; Xu, W.; et al. Lipid Nanoparticle Delivery of Chemically Modified NGF(R100W) mRNA Alleviates Peripheral Neuropathy. Adv. Healthc. Mater. 2023, 12, e2202127.

57. Guo, C.; Ma, X.; Gao, F.; Guo, Y. Off-Target Effects in CRISPR/Cas9 Gene Editing. Front. Bioeng. Biotechnol. 2023, 11, 1143157.
58. Evers, M.J.W.; Du, W.; Yang, Q.; Kooijmans, S.A.A.; Vink, A.; van Steenbergen, M.; Vader, P.; de Jager, S.C.A.; Fuchs, S.A.; Mastrobattista, E.; et al. Delivery of Modified mRNA to Damaged Myocardium by Systemic Administration of Lipid Nanoparticles. J. Control. Release 2022, 343, 207–216.
59. Chen, F.; Alphonse, M.; Liu, Q. Strategies for Nonviral Nanoparticle-Based Delivery of CRISPR/Cas9 Therapeutics. Wiley Interdiscip. Rev. Nanomed. Nanobiotechnol. 2020, 12, e1609.
60. Greig, J.A.; Martins, K.M.; Breton, C.; Lamontagne, R.J.; Zhu, Y.; He, Z.; White, J.; Zhu, J.X.; Chichester, J.A.; Zheng, Q.; et al. Integrated Vector Genomes May Contribute to Long-Term Expression in Primate Liver after AAV Administration. Nat. Biotechnol. 2023.
61. Nathwani, A.C.; Reiss, U.M.; Tuddenham, E.G.; Rosales, C.; Chowdary, P.; McIntosh, J.; Della Peruta, M.; Lheriteau, E.; Patel, N.; Raj, D.; et al. Long-Term Safety and Efficacy of Factor IX Gene Therapy in Hemophilia B. N. Engl. J. Med. 2014, 371, 1994–2004.
62. Hanlon, K.S.; Kleinstiver, B.P.; Garcia, S.P.; Zaborowski, M.P.; Volak, A.; Spirig, S.E.; Muller, A.; Sousa, A.A.; Tsai, S.Q.; Bengtsson, N.E.; et al. High Levels of AAV Vector Integration into CRISPR-Induced DNA breaks. Nat. Commun. 2019, 10, 4439.
63. Kim, D.; Luk, K.; Wolfe, S.A.; Kim, J.S. Evaluating and Enhancing Target Specificity of Gene-Editing Nucleases and Deaminases. Annu. Rev. Biochem. 2019, 88, 191–220.
64. Ghani, M.W.; Iqbal, A.; Ghani, H.; Bibi, S.; Wang, Z.; Pei, R. Recent Advances in Nanocomposite-Based Delivery Systems for Targeted CRISPR/Cas Delivery and Therapeutic Genetic Manipulation. J. Mater. Chem. B 2023, 11, 5251–5271.
65. Kulishova, L.M.; Vokhtantsev, I.P.; Kim, D.V.; Zharkov, D.O. Mechanisms of the Specificity of the CRISPR/Cas9 System in Genome Editing. Mol. Biol. 2023, 57, 269–284.
66. Ran, F.A.; Hsu, P.D.; Lin, C.Y.; Gootenberg, J.S.; Konermann, S.; Trevino, A.E.; Scott, D.A.; Inoue, A.; Matoba, S.; Zhang, Y.; et al. Double Nicking by RNA-Guided CRISPR Cas9 for Enhanced Genome Editing Specificity. Cell 2013, 154, 1380–1389.
67. Slaymaker, I.M.; Gao, L.; Zetsche, B.; Scott, D.A.; Yan, W.X.; Zhang, F. Rationally Engineered Cas9 Nucleases with Improved Specificity. Science 2016, 351, 84–88.
68. Kleinstiver, B.P.; Pattanayak, V.; Prew, M.S.; Tsai, S.Q.; Nguyen, N.T.; Zheng, Z.; Joung, J.K. High-Fidelity CRISPR-Cas9 Nucleases with No Detectable Genome-Wide off-Target Effects. Nature 2016, 529, 490–495.
69. Chen, J.S.; Dagdas, Y.S.; Kleinstiver, B.P.; Welch, M.M.; Sousa, A.A.; Harrington, L.B.; Sternberg, S.H.; Joung, J.K.; Yildiz, A.; Doudna, J.A. Enhanced Proofreading Governs CRISPR-Cas9

Targeting Accuracy. Nature 2017, 550, 407–410.

70. Bravo, J.P.K.; Liu, M.S.; Hibshman, G.N.; Dangerfield, T.L.; Jung, K.; McCool, R.S.; Johnson, K.A.; Taylor, D.W. Structural Basis for Mismatch Surveillance by CRISPR-Cas9. Nature 2022, 603, 343–347.
71. Casini, A.; Olivieri, M.; Petris, G.; Montagna, C.; Reginato, G.; Maule, G.; Lorenzin, F.; Prandi, D.; Romanel, A.; Demichelis, F.; et al. A Highly Specific SpCas9 Variant is Identified by In Vivo Screening in Yeast. Nat. Biotechnol. 2018, 36, 265–271.
72. Lee, J.K.; Jeong, E.; Lee, J.; Jung, M.; Shin, E.; Kim, Y.H.; Lee, K.; Jung, I.; Kim, D.; Kim, S.; et al. Directed Evolution of CRISPR-Cas9 to Increase Its Specificity. Nat. Commun. 2018, 9, 3048.
73. Vakulskas, C.A.; Dever, D.P.; Rettig, G.R.; Turk, R.; Jacobi, A.M.; Collingwood, M.A.; Bode, N.M.; McNeill, M.S.; Yan, S.; Camarena, J.; et al. A High-Fidelity Cas9 Mutant Delivered as a Ribonucleoprotein Complex Enables Efficient Gene Editing in Human Hematopoietic Stem and Progenitor Cells. Nat. Med. 2018, 24, 1216–1224.
74. Hu, J.H.; Miller, S.M.; Geurts, M.H.; Tang, W.; Chen, L.; Sun, N.; Zeina, C.M.; Gao, X.; Rees, H.A.; Lin, Z.; et al. Evolved Cas9 Variants with Broad PAM Compatibility and High DNA Specificity. Nature 2018, 556, 57–63.
75. Pattanayak, V.; Lin, S.; Guilinger, J.P.; Ma, E.; Doudna, J.A.; Liu, D.R. High-throughput Profiling of off-Target DNA Cleavage Reveals RNA-Programmed Cas9 Nuclease Specificity. Nat. Biotechnol. 2013, 31, 839–843.
76. Cong, L.; Ran, F.A.; Cox, D.; Lin, S.; Barretto, R.; Habib, N.; Hsu, P.D.; Wu, X.; Jiang, W.; Marraffini, L.A.; et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. Science 2013, 339, 819–823.
77. Anders, C.; Niewoehner, O.; Duerst, A.; Jinek, M. Structural Basis of PAM-Dependent Target DNA Recognition by the Cas9 Endonuclease. Nature 2014, 513, 569–573.
78. Cho, S.W.; Kim, S.; Kim, Y.; Kweon, J.; Kim, H.S.; Bae, S.; Kim, J.S. Analysis of off-Target Effects of CRISPR/Cas-Derived RNA-Guided Endonucleases and Nickases. Genome Res. 2014, 24, 132–141.
79. Kawamata, M.; Suzuki, H.I.; Kimura, R.; Suzuki, A. Optimization of Cas9 Activity through the Addition of Cytosine Extensions to Single-Guide RNAs. Nat. Biomed. Eng. 2023, 7, 672–691.
80. Wyvekens, N.; Topkar, V.V.; Khayter, C.; Joung, J.K.; Tsai, S.Q. Dimeric CRISPR RNA-Guided FokI-dCas9 Nucleases Directed by Truncated gRNAs for Highly Specific Genome Editing. Hum. Gene Ther. 2015, 26, 425–431.
81. Rozners, E. Chemical Modifications of CRISPR RNAs to Improve Gene-Editing Activity and Specificity. J. Am. Chem. Soc. 2022, 144, 12584–12594.