

Modulating CRISPR-Cas genome editing using guide-complementary DNA oligonucleotides

Thomas Swartjes^{1,*}, Peng Shang^{2,*}, Dennis van den Berg³, Tim A. Künne^{1,4}, Niels Geijsen², Stan J.J. Brouns^{4,5}, John van der Oost¹, Raymond H.J. Staals^{1,**}, Richard A. Notebaart^{3,**}

1. Laboratory of Microbiology, Wageningen University & Research, Wageningen, The Netherlands
2. Department of Anatomy and Embryology, Leiden University Medical Centre, Leiden, The Netherlands
3. Food Microbiology, Wageningen University & Research, Wageningen, The Netherlands
4. Department of Bionanoscience, Delft University of Technology, The Netherlands
5. Kavli Institute of Nanoscience, Delft, The Netherlands

* authors contributed equally

** corresponding authors (raymond.staals@wur.nl and richard.notebaart@wur.nl)

Abstract

CRISPR-Cas has revolutionized genome editing and has a great potential for applications, such as correcting human genetic disorders. To increase the safety of genome editing applications, CRISPR-Cas may benefit from strict control over Cas enzyme activity. Previously, anti-CRISPR proteins and designed oligonucleotides have been proposed to modulate CRISPR-Cas activity. Here we report on the potential of guide-complementary DNA oligonucleotides as controlled inhibitors of Cas9 ribonucleoprotein complexes. First, we show that DNA oligonucleotides down-regulate Cas9 activity in human cells, reducing both on and off-target cleavage. We then used *in vitro* assays to better understand how inhibition is achieved and under which conditions. Two factors were found to be important for robust inhibition: the length of the complementary region, and the presence of a PAM-loop on the inhibitor. We conclude that DNA oligonucleotides can be used to effectively inhibit Cas9 activity both *ex vivo* and *in vitro*.

Introduction

CRISPR-Cas (clustered regularly interspaced short palindromic repeats and CRISPR-associated proteins) systems provide prokaryotes with adaptive immunity against mobile genetic elements (MGEs)^{1,2}. Similar to other CRISPR-Cas systems, the Cas9 nuclease from *Streptococcus pyogenes* (SpCas9) mediates double stranded cleavage of the target DNA^{3,4}. In brief, the Cas protein binds a guide RNA (gRNA) to form a ribonucleoprotein (RNP) complex. This complex then interrogates the DNA to find a DNA sequence complementary to the gRNA (protospacer)⁵. To this end, the RNP complex binds DNA sequences with a protospacer adjacent motif (PAM) and causes initial unwinding of the adjacent DNA bases³⁻⁸. If these bases are complementary to the crRNA, DNA unwinding proceeds along the gRNA until a stable R-loop is formed^{5,9}. Lastly, both DNA strands are cleaved, resulting in a double-strand break (DSB) in the target DNA^{3,4}.

Cas9 and other Cas nucleases are used for genome editing by introducing DSBs at specific DNA sequences. One possibility is that the genomic edits occur independently of the Cas nuclease. The nuclease would then be used as a counter-selection system to select against non-edited versions of the target site by introducing DSBs ¹⁰. Alternatively, the Cas nuclease might first introduce a DSB, which then induces local DNA repair ¹¹. The two most common repair mechanisms that act on DSBs are homology directed repair (HDR) and non-homologous end joining (NHEJ) ¹². HDR uses a homologous repair template to fix the DSB according to the template sequence ¹². NHEJ resolves the double-stranded break without the need for a repair template, which often results in insertions or deletions (indels) at the site of the DSB formation ¹³.

For genome editing applications of CRISPR-Cas, it is crucial that the Cas nuclease introduces DSBs at the sequence of interest with sufficient efficacy. However, Cas-nucleases were found to also create DSBs at sequences with imperfect complementarity to the gRNA ^{14–18}. It is essential to prevent genetic changes at such off-target sites, especially for therapeutic genome editing. In addition, strict control over Cas9 activity might be used to confine DSB formation to the desired cells in a limited time-frame.

In nature, Cas enzyme activity can be inhibited by phage-encoded anti-CRISPR (ACR) proteins that act on different stages of CRISPR-Cas based immunity ^{19–21}. The ACRs that were found to inhibit Cas nuclease activity, have been reported to be useful for controlling CRISPR-Cas-based genome editing ^{21–26}. Aside from these naturally occurring ACRs, several other strategies have been devised to control Cas enzyme activity at the level of transcription ^{27–29}, translation ^{30–32}, protein state ^{33–47}, and guide RNA ^{14,48}.

In addition, single stranded DNA or RNA molecules can be designed to inhibit Cas nucleases. Such oligonucleotide-based inhibitors provide several advantages compared to natural ACR proteins. DNA oligos are inexpensive, can be rapidly manufactured, and could provide a

systematic way to inhibit different Cas nucleases, whereas the use of anti-CRISPRs is dependent on compatibility with the Cas nuclease of choice.

Oligo-based inhibitors have been shown to work *in vitro* and in human cell cultures for Cas9 and Cas12a⁴⁹⁻⁵¹. Potent inhibition of Cas9 was observed with RNA-DNA hybrids or chemically-modified DNA inhibitors which interact with the repeat sequence of the guide RNA or with the PAM-interacting domain of Cas9⁵⁰. In addition, truncated gRNA designs have been shown to allow dsDNA binding, but not cleavage by Cas9^{52,53}. Such truncated gRNAs were used to specifically direct non-cleaving Cas9 RNPs to off-target sequences, thereby preventing active Cas9 RNP from binding these off-target sites⁵¹. Lastly, DNA oligos with phosphorothioate linkages displayed strong inhibition of Cas12a activity, apparently independent of the nucleotide sequence of the inhibitor used⁴⁹.

In the current study, we assessed whether SpCas9 could be inhibited with guide-complementary DNA oligonucleotides without any chemical modifications. To this end, we designed and tested different oligo-based inhibitors complementary to the spacer-derived part of the guide RNA (these oligos thus have the same sequence as the PAM-proximal part of the protospacer). We also investigated the effect of extending the inhibitors with a double-stranded PAM sequence. We show that various designs provide strong, sequence-dependent inhibition of Cas9 *in vitro* and *ex vivo*.

In addition, we investigated the effect of oligo-based Cas9 inhibition on off-target sites in the context of genome editing. We found that while inhibition reduces both on- and off-target activity of Cas9, the presence of specific oligo designs results in slightly increased specificity. By comparing the inhibitor results with a Cas9 titration, we concluded that the increased specificity is a general consequence of lowering overall Cas9 activity.

Lastly, we studied which mechanisms lead to the observed inhibition. We concluded that the effect of the tested inhibitors is dependent both on the length of the oligo-based inhibitors and

on the presence of a PAM-loop. Their relative importance is strongly affected by the speed at which Cas9 cleaves the targeted DNA.

Results

To assess whether guide-complementary oligonucleotides could work as guide-specific inhibitors of Cas nucleases, we designed guide-complementary DNA oligos (hereafter referred to as ‘inhibitors’) of 8 and 20 nucleotides (nt) (Fig. 1). The inhibitors are complementary to the PAM-proximal, spacer-derived part of the guide RNA for SpCas9 (Fig. 1B,C). We also designed oligos with a 5'-extension intended to loop and fold back onto itself, creating a double-stranded PAM (Fig. 1D).

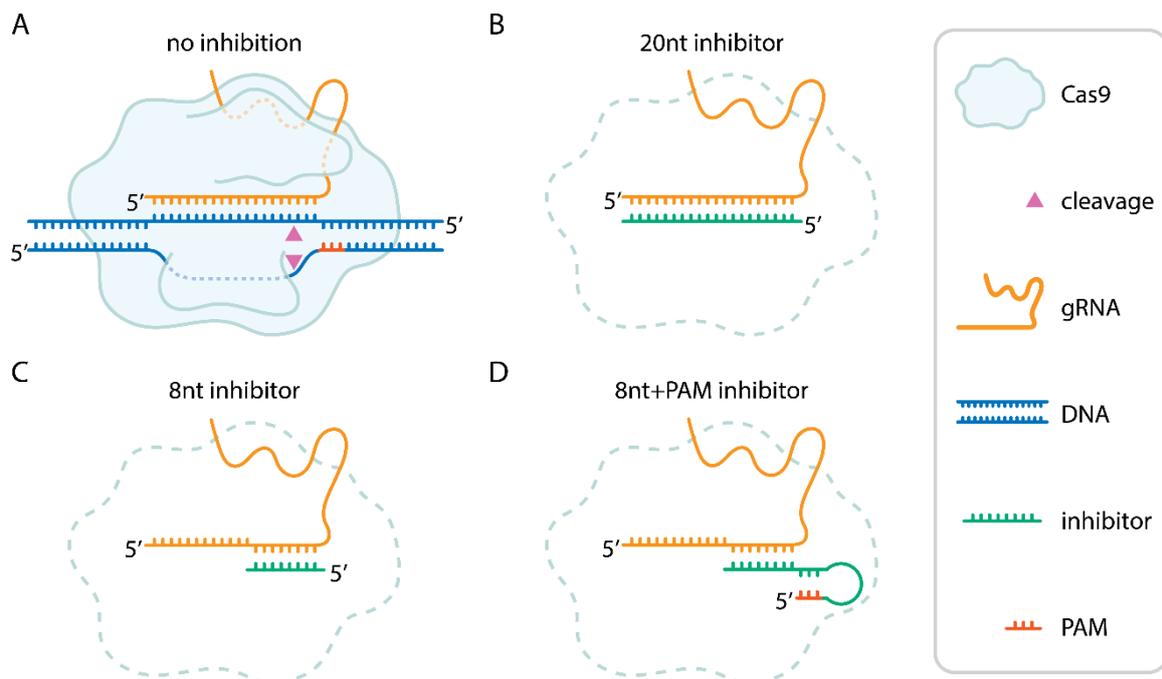


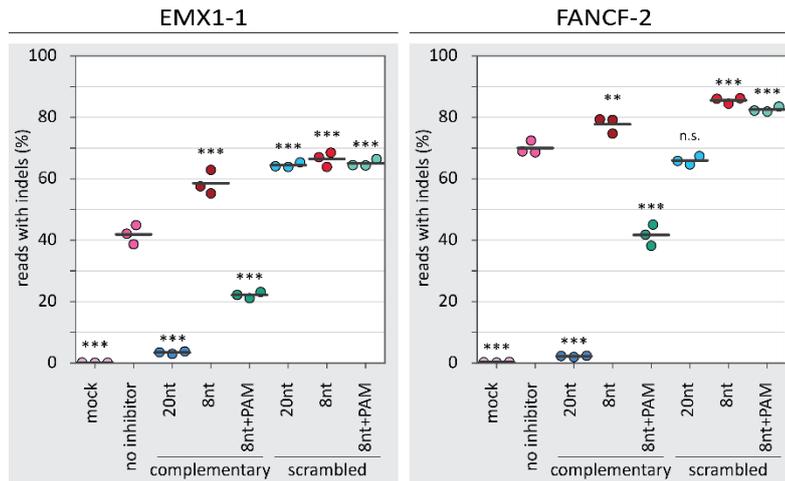
Fig. 1. Design of guide-complementary oligo-based inhibitors.

Schematic representation of a Cas9 nuclease, guide RNA (here displayed as a single guide RNA, sgRNA) with and without oligo-based inhibitors. (A) without inhibitors, the RNP binds the target sequence in genomic DNA, forms an R-loop and proceeds to cleave both strands of the DNA (triangles). (B) 20nt inhibitor. (C) 8nt inhibitor. (D) 8nt+PAM inhibitor. For all nucleic acids, 5' ends are indicated.

We set out to test the effects of these DNA oligo-based inhibitors in a genome editing context. To that end, we delivered purified SpCas9 protein, guide RNA, and the inhibitors to Chronic myelogenous leukemia (CML) cells using iTOP⁵⁴. We analyzed editing of two endogenous gene loci, EMX1-1 and FANCF-2, under all test conditions. iTOP-transduced cells were allowed to grow to 90% of confluency, we then extracted the genomic DNA from the cells. From the genomic DNA, we amplified six regions of interest: the two on target sites (EMX1-1 and FANCF-2) and two, previously identified^{55,56}, prominent off-target sites for each (Supplementary Table 1). We performed deep-sequencing on the amplicons to assess the frequency of indels arisen from NHEJ.

Without inhibitors, iTOP-delivered Cas9 RNP achieved roughly 40% and 70% on-target indel frequencies on EMX1-1 and FANCF-2 loci, respectively (Fig. 2A). Addition of the 20nt inhibitor - in equimolar ratio to the Cas9 RNP - strongly reduces indel frequency (to less than 5%), indicating that Cas9-induced DSB formation is repressed. Repression is also observed with the 8nt+PAM inhibitor, but to a lesser extent (to between 20% and 50%). The 8nt inhibitor however, unexpectedly showed increased indel frequencies rather than inhibition. The same was observed for most of the sequence-scrambled versions of the inhibitors, which were included as a control for sequence specificity. Together, this shows that both the 20nt and 8nt+PAM inhibitors repress Cas9 activity *ex vivo* in a sequence-specific manner.

A



B

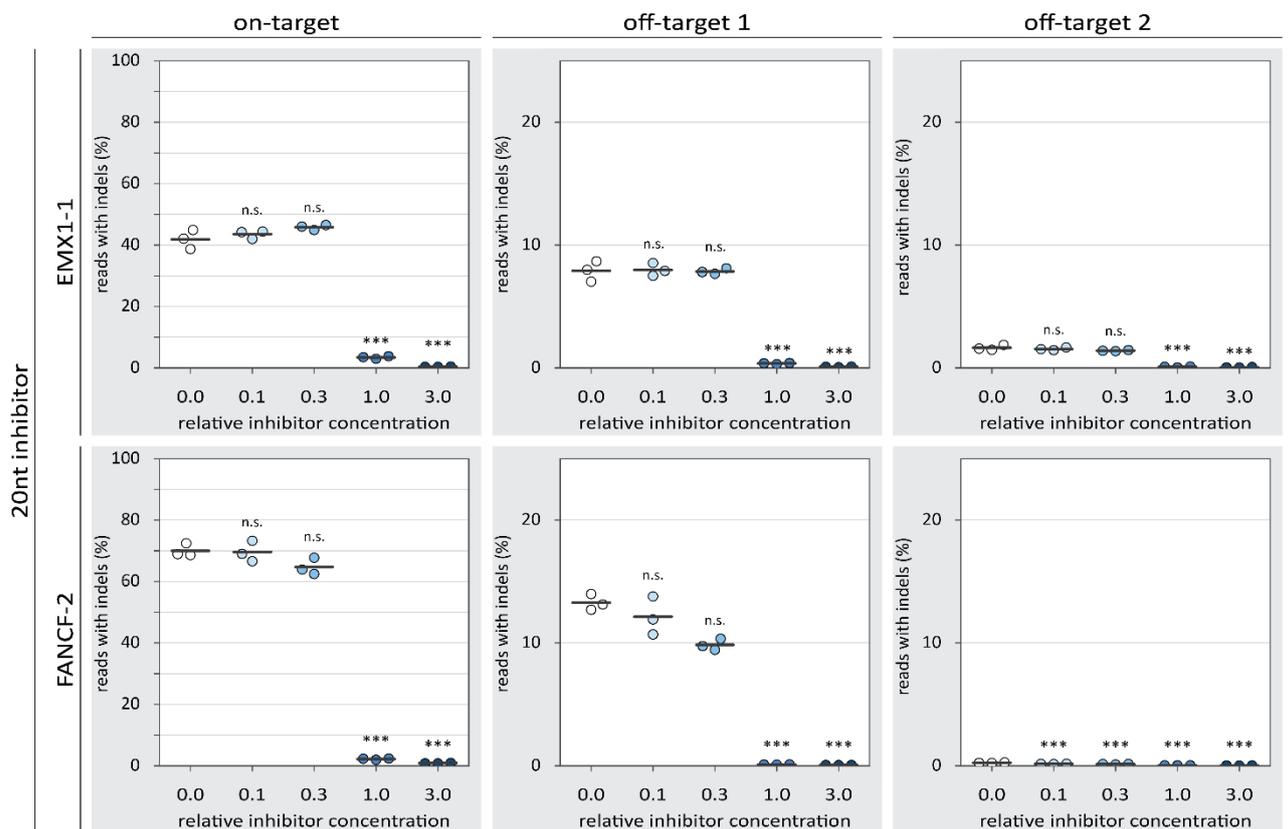


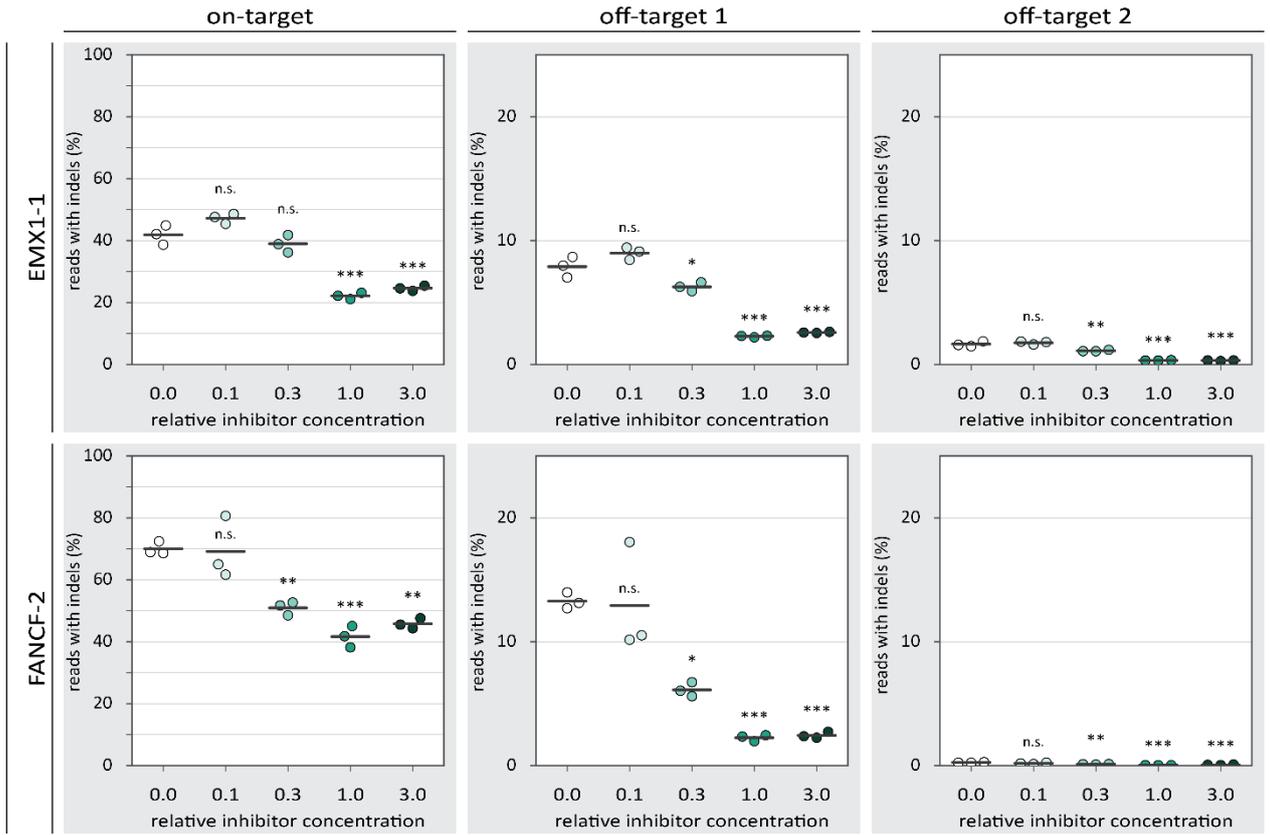
Fig. 2. inhibition of Cas9 by guide-complementary DNA oligos.

Individual replicates are displayed as colored dots while the horizontal black lines show the mean of the three replicates. *p*-values were calculated with a one-way ANOVA and subsequent Tukey's test. 'n.s.': not significant; *: *p*-value < 0.05; **: *p*-value < 0.01; *** *p*-value < 0.001. **A:** Percentage of reads containing indels for on-target loci of EMX1-1 and FANCF-2. The 'mock' condition constitutes a control where the iTOP method was used without delivering Cas9, guide RNA and inhibitor. All conditions except the 'mock' used SpCas9, guide RNA and – except for 'no inhibitor' – a DNA oligo design as detailed in Fig. 1. The DNA oligos were delivered at a molar concentration equal to the concentration of Cas9 and guide RNA. *p*-values are displayed for the 'mock' condition compared to each other condition. **B:** Percentage of reads containing indels for on- and off-target loci of EMX1-1 and FANCF-2 for different concentrations of the 20nt inhibitor. The 'relative inhibitor concentration' indicates the molar concentration of the inhibitor relative to Cas9 and guide RNA. A 'relative inhibitor concentration' of 1.0 means that the used molar concentration of the inhibitor is equal to that of Cas9 and guide RNA. For off-target loci, the y-axis is adjusted because relatively few indels contained indels in these conditions. *p*-values are displayed for the 0.0 relative oligo concentration compared to each other condition.

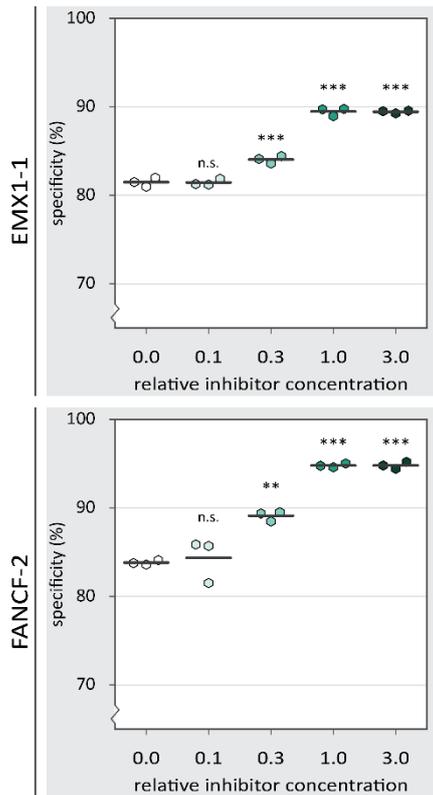
To characterize the observed inhibition in more detail, we tested the inhibitors in a concentration gradient. Substantial inhibition by the 20nt inhibitor is only observed at and above equimolar ratio to the Cas9 RNP (Fig. 2B). For the 8nt+PAM inhibitors, a more gradual dose-response is observed (Fig. 3A) with low concentrations seemingly providing low levels of inhibition. However, even at the highest concentration, the 8nt+PAM inhibitors are not able to reduce indel frequencies to the same extent as was achieved with the 20nt inhibitors. At the investigated off-target loci, we observed reduced indel frequencies with the 20nt and 8nt+PAM inhibitors, following a similar dose-response as was seen on-target.

The 8nt+PAM oligo displays inhibition both on- and off-target, but neither is completely abolished. This raises the question to what extent the ratio between on- and off-targeting is affected by the inhibitors. We calculated a specificity score by dividing the on-target indel frequency by the sum of the on- and off-target indel frequencies. This specificity score signifies the percentage of indels that were mapped to the on-target locus, corrected for the total number of reads mapped to each locus. We found that the higher concentrations of the 8nt+PAM oligo slightly increase the specificity score (Fig. 3B). However, it is important to realize that at the same inhibitor concentrations, on-target activity is substantially reduced (Fig. 3B,C).

A



B



C

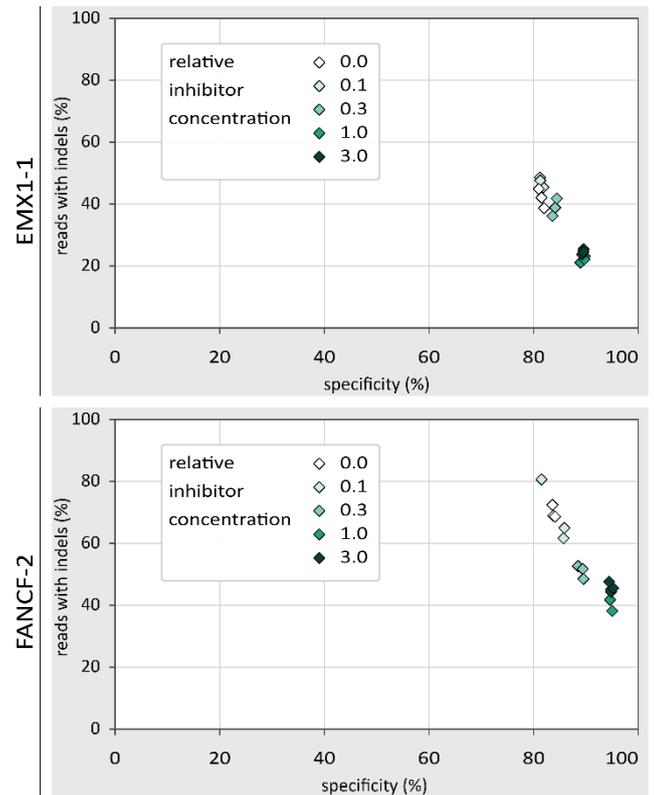
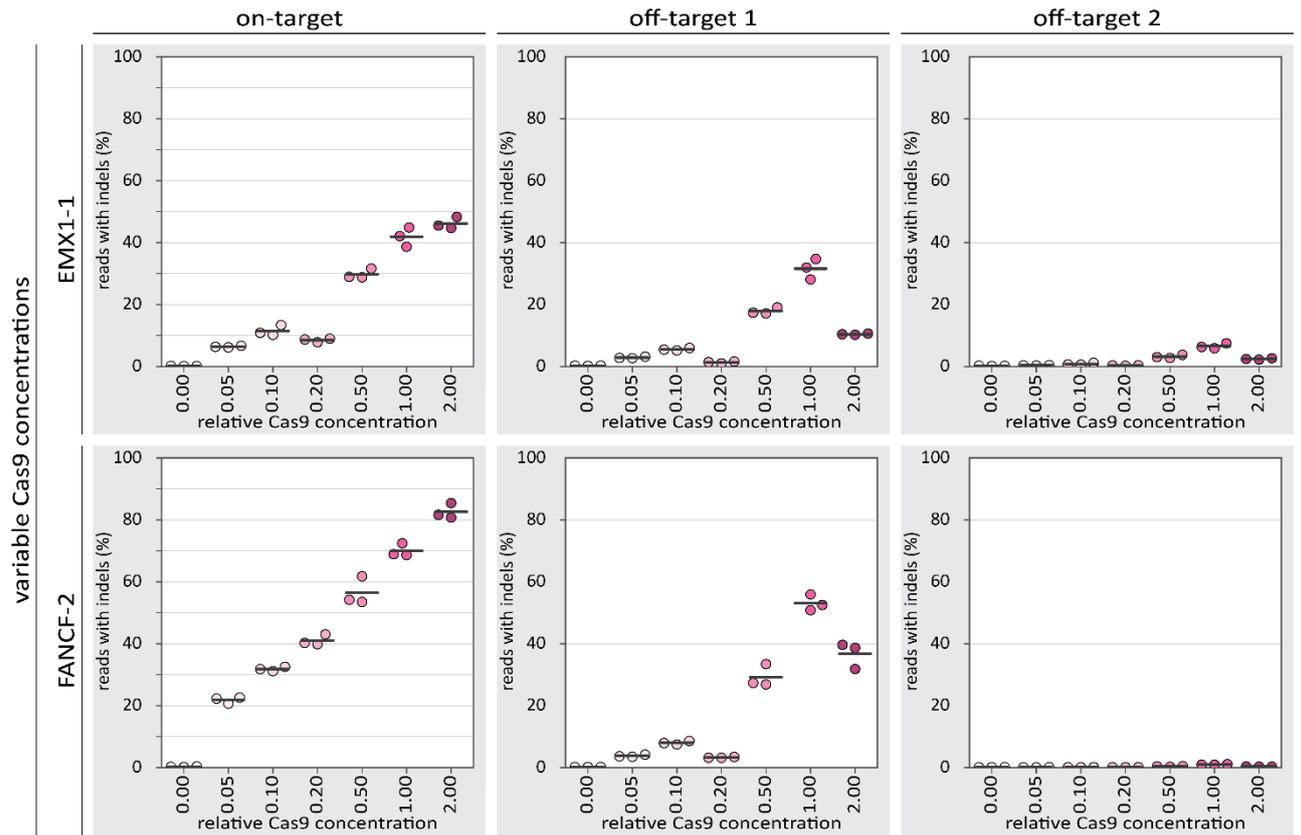


Fig. 3. The effects of 8nt+PAM inhibitor concentration on specificity of Cas9-mediated NHEJ.

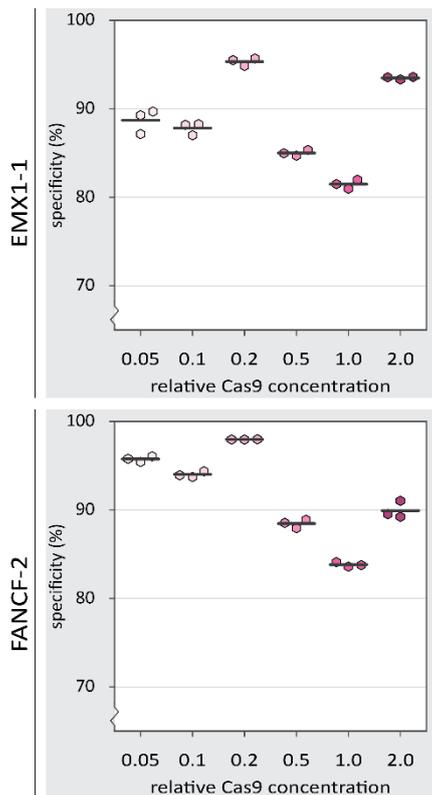
p-values were calculated with a one-way ANOVA and subsequent Tukey's test. *p*-values are displayed for the 0.0 relative oligo concentration compared to each other condition. 'n.s.': not significant; *: *p*-value < 0.05; **: *p*-value < 0.01; *** *p*-value < 0.001. **A:** Percentage of reads containing indels for on- and off-target loci of EMX1-1 and FANCF-2 for different concentrations of the 8nt+PAM inhibitor. The 'relative inhibitor concentration' indicates the molar concentration of the inhibitor relative to Cas9 and guide RNA. A 'relative inhibitor concentration' of 1.0 means that the used molar concentration of the inhibitor is equal to that of Cas9 and guide RNA. For off-target loci, the y-axis is adjusted because relatively few indels contained indels in these conditions. **B:** Percentage specificity for on-target loci of EMX1-1 and FANCF-2 for different concentrations of the 8nt+PAM inhibitor. The y-axis starts at 70% specificity to better visualize the differences between conditions. The individual replicates are displayed as colored hexagons while the horizontal black lines show the mean of the three replicates. **C:** comparison of *percentage reads with indels* and *percentage specificity* for different concentrations of the 8nt+PAM inhibitor. For each inhibitor concentration, individual replicates are displayed as diamonds with the same color.

We next asked whether the observed increase in specificity may be the result of a decreased subpopulation of active Cas9 RNP complexes through inhibition with DNA oligos. To address this, we looked at a range of Cas9 RNP concentrations to see how they affect on-target activity and specificity. At both on-target loci, we found that increasing RNP concentrations lead to higher indel frequencies (Fig. 4A). Strikingly, at the off-target sites, the highest Cas9 RNP concentration yielded relatively low indel frequencies. The resulting specificity scores are highest at relatively low RNP concentrations, and drop to a minimum at high concentrations. At the highest concentration however, specificity increases again, due to the unexpected decrease in off-target indels at that RNP concentration (Fig. 4B). When plotting both the on-target activity and specificity, among the intermediate Cas9 concentrations we observed a similar trade-off between activity and specificity as was observed with the 8nt+PAM inhibitor (Fig. 4C). The lowest and highest Cas9 RNP concentration do not follow this pattern however.

A



B



C

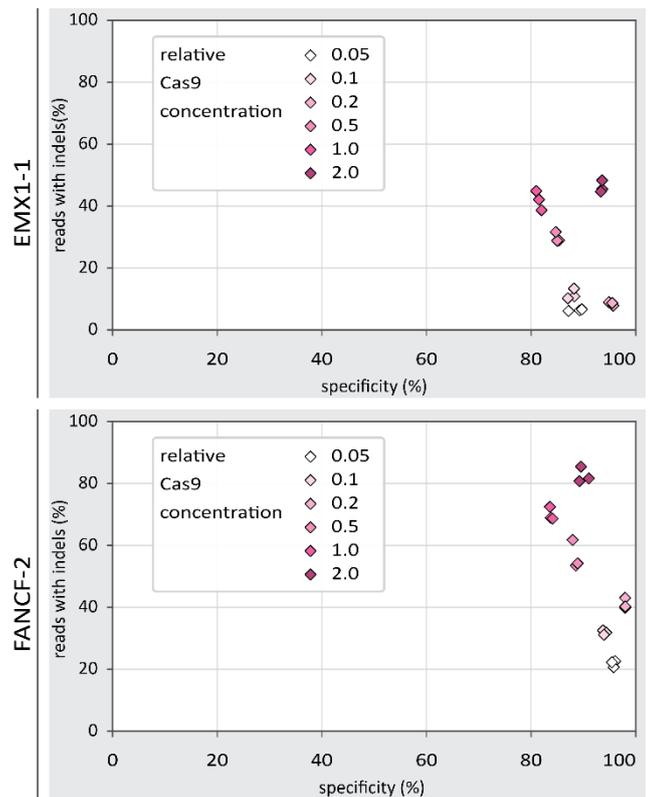


Fig. 4. The effects of Cas9 concentration on specificity of Cas9-mediated NHEJ.

A: Percentage of reads containing indels for on- and off-target loci of EMX1-1 and FANCF-2 for different concentrations of SpCas9 RNP in the absence of DNA oligo-based inhibitors. The individual replicates are displayed as colored dots while the horizontal black lines show the mean of the three replicates. The concentration 0.00 is the 'mock' condition in Fig. 2A. **B:** Percentage specificity for on-target loci of EMX1-1 and FANCF-2 for different concentrations of Cas9 RNP. The y-axis starts at 70% specificity to better visualize the differences between conditions. The individual replicates are displayed as colored hexagons while the horizontal black lines show the mean of the three replicates. The 0.00 Cas9 RNP concentration is not included here because the specificity values make no sense if there are virtually no indels at all. **C:** comparison of *percentage reads with indels* and *percentage specificity* for different concentrations of the Cas9 RNP. For each inhibitor concentration, individual replicates are displayed as diamonds with the same color. Again, the 0.00 Cas9 RNP concentration is not included here.

So far, we have seen that the 20nt inhibitor provides the most potent inhibition in our *ex vivo* setup, almost abolishing on- and off-target indel formation. The 8nt+PAM version displayed less inhibition, but shows some level of inhibition already at low concentrations. Next, we designed an experimental approach to better understand the mechanism by which the inhibitors function and shed light on why the 20nt and 8nt+PAM inhibitors display different dose-response *ex vivo*. To investigate the underlying mechanism, we used *in vitro* cleavage assays, which provide more experimental control than the *ex vivo* setup.

In preparation for the *in vitro* experiments, we performed PCRs of the EMX1-1 and FANCF-2 loci to produce the cleavage substrate DNA (supplementary table 4). Firstly, we incubated the purified Cas9 protein with the appropriate guide RNA (20 minutes pre-incubation at 25°C: Cas9 + gRNA). Secondly, we added the target dsDNA and the oligo-based inhibitors to the Cas9 and guide and incubated to allow cleavage of the DNA (30 minutes 37°C). We then used agarose gel electrophoresis to assess how much of the substrate DNA had been cleaved. We found that the pre-incubated Cas9 and guide RNA could efficiently digest the substrate DNA (Fig. 5). However, addition of the 8nt+PAM inhibitor substantially reduced DNA cleavage. In agreement with the aforementioned *ex vivo* analyses, we did not observe an effect from the 8nt inhibitor, nor with the sequence scrambled inhibitors. Unexpectedly, the 20nt inhibitor also did not show a clear reduction in DNA cleavage, despite strongly inhibiting indel formation *ex vivo*.

To better understand the differences between *ex vivo* and *in vitro* outcomes we repeated the *in vitro* cleavage assays, but this time without the initial incubation of Cas9 with guide RNA

(pre-incubation: none). In addition, we now included a 20nt+PAM inhibitor design in both experimental setups to see whether the PAM loop could rescue the performance of the 20nt inhibitor.

We now found that the 20nt inhibitor strongly inhibited DNA cleavage (Fig. 5). The 20nt+PAM inhibitor displayed strong inhibition in both experimental setups. This shows that inhibition without a PAM-loop is dependent on the order in which inhibitor, Cas9 and guide RNA are mixed. In contrast, the PAM-loop enables inhibition regardless of the mixing order.

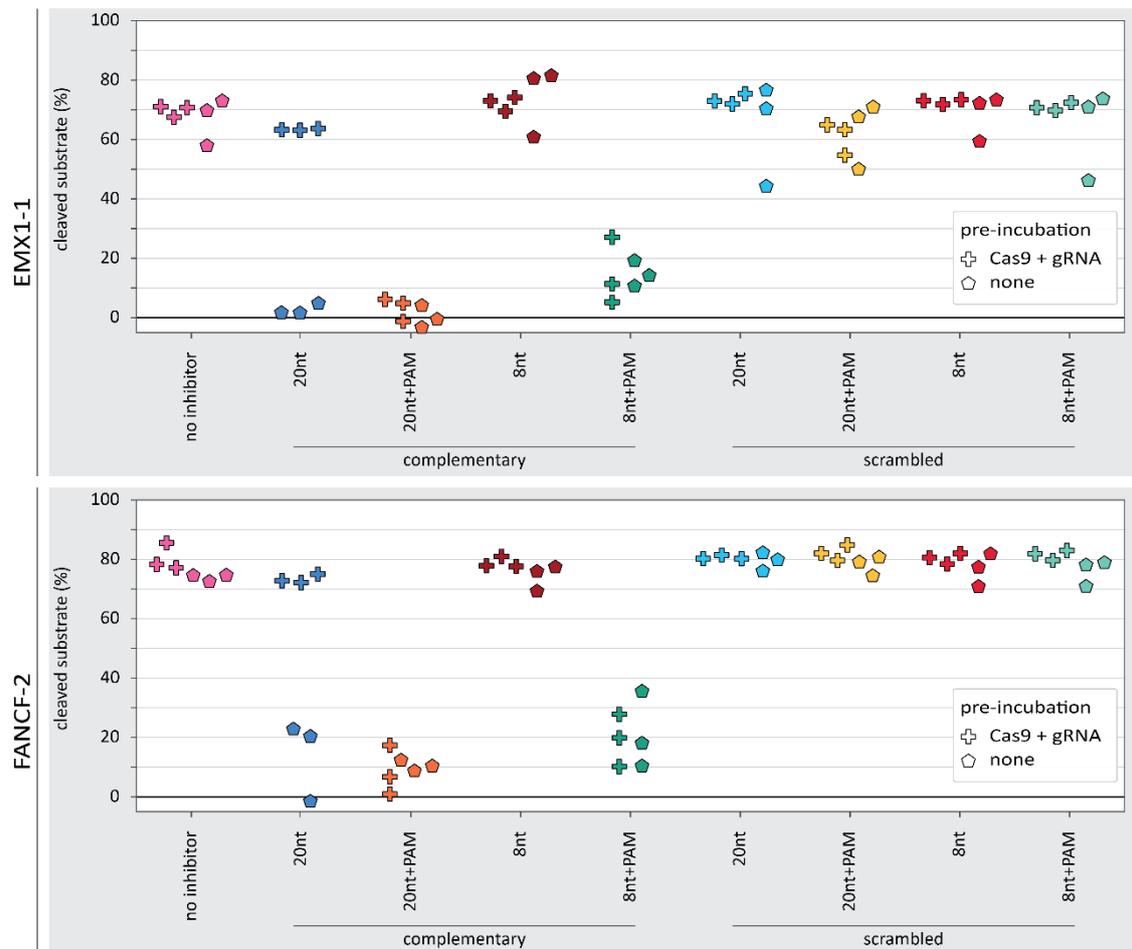


Fig. 5. *in vitro* DNA cleavage in the presence of various oligo-based inhibitor designs.

The percentage of DNA substrate cleaved as measured from band intensity on agarose gel. The plus-markers show the observed percentage of DNA cleavage when Cas9 and guide RNA were pre-incubated. The pentagon-markers represent replicates without pre-incubation. For EMX1-1 with oligo & gRNA pre-incubation, the scrambled oligo design each have a single datapoint that displays relatively poor cleavage. These low datapoints are all derived from a single batch of substrate DNA purified from agarose gel, different from the other datapoints.

We wondered whether the PAM-less 20nt inhibitor would only be able to bind the guide RNA when the RNA is not yet bound by Cas9. To assess this, we performed an electrophoretic mobility shift assay (EMSA) to visualize binding between inhibitor, Cas9 and guide RNA in different combinations. From this, we observed a large shift upon adding Cas9 to the guide RNA, indicating that the two bind to each other (Fig. 6A). If the 20nt inhibitor is added as well, the bands shift up slightly more, regardless of the order in which the components are mixed. This is not the case with the scrambled 20nt inhibitors, indicating that binding of the inhibitor is sequence-specific. Overall, we conclude that the 20nt inhibitor can bind both pre-formed RNP complexes and free guide RNA. It is therefore likely that the 20nt inhibitor can also inhibit at both stages.

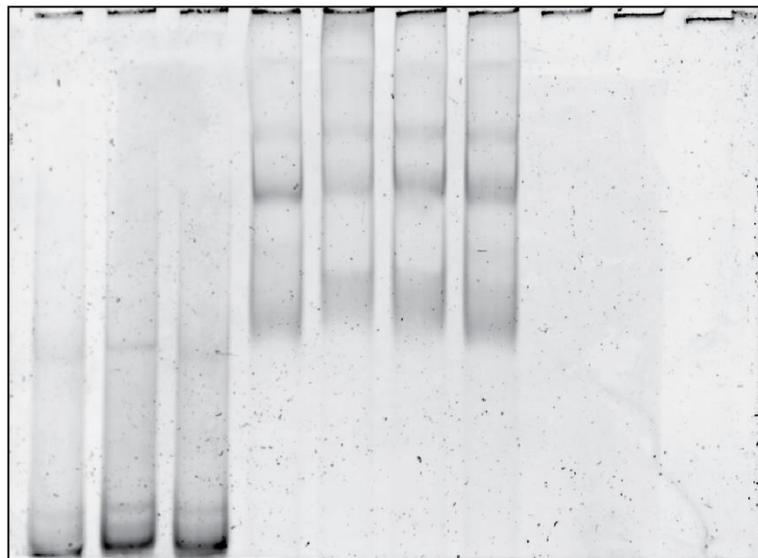
We then asked why – given the EMSA results – we did not observe inhibition with the 20nt inhibitor on pre-incubated Cas9 and guide RNA (Fig. 5). We reasoned that the 20nt inhibitor might require a longer time to bind pre-formed RNP compared to free guide RNA. If that is the case, the 20nt inhibitor would simply not have had enough time to bind (thus inhibiting cleavage of the DNA substrate by Cas9 RNP). We then tested how much time the 20nt inhibitor requires for inhibition of pre-formed Cas9 RNP. To that end, we first incubated the guide RNA and Cas9 to form the RNP. To this pre-formed RNP, we then incubated the 20nt or 8nt+PAM oligo for a variable duration before adding the substrate DNA and allowing 30 minutes of digestion.

We observed that the 20nt inhibitor requires roughly one minute to prevent most of the DNA cleavage upon subsequent addition of the substrate DNA (Fig. 6B). Inhibition by the 8nt+PAM design appears to be instantaneous. This explains why we did not observe a substantial reduction in DNA cleavage when the 20nt inhibitor was added simultaneously with the substrate DNA. Taken together, the *in vitro* results presented here show that a PAM-loop affects how the oligo-based inhibitors function. Designs with a PAM-loop can efficiently inhibit DNA cleavage regardless of whether they are added to free guide RNA, or to pre-formed RNP

complexes. In contrast, PAM-less designs are effective when added to free guide RNA, and require a longer incubation to inhibit pre-formed RNP complexes.

A

Cas9	-	-	-	+	2	1	1	+	+	+
guide RNA	+	+	+	+	1	1	1	-	-	-
20nt inhibitor	-	+	-	-	1	2	-	+	-	-
scrambled 20nt	-	-	+	-	-	-	2	-	+	-



B

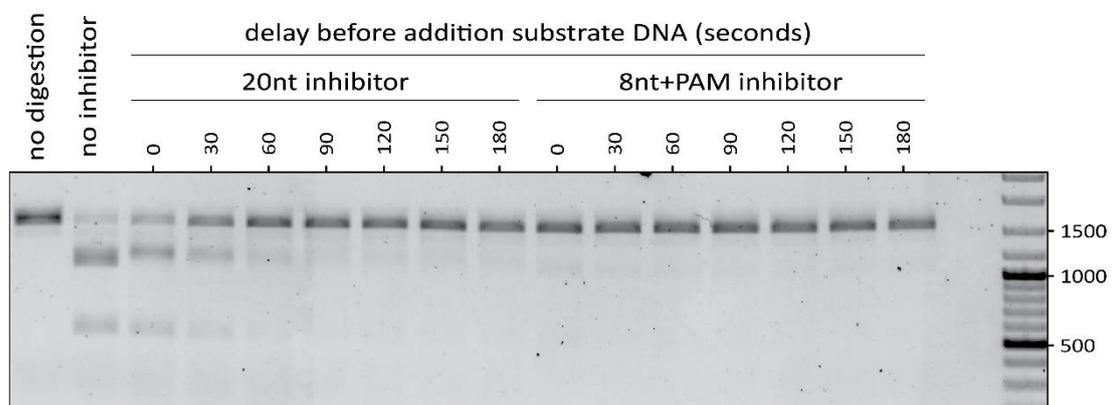


Fig. 6. Timing-dependent inhibition of Cas9 by oligo-based inhibitors.

A: electrophoretic mobility shift assay on 5% polyacrylamide gel, stained with SYBR Gold for nucleic acids visualization. Components with a '+' sign were present, whereas those with a '-' sign were not. A '1' indicates that these components were initially mixed. In contrast, the components marked with '2' were added after 15 min of initial incubation. The black lines at the top of the image correspond to the bottom of the wells of the gel. **B:** time-lapse assay where the 20nt or 8nt+PAM inhibitors were added to pre-formed RNP complexes, then incubated for variable durations ('delays') before addition of the substrate DNA. The Generuler mix ladder was included on the far right and relevant fragment lengths are indicated in base-pairs. The linear substrate DNA is 1500bp long and cleavage by Cas9 would result in 2 fragments of lengths 1000bp and 500bp.

Discussion & Conclusion

We have described sequence-dependent inhibition of Cas9 with guide-complementary unmodified DNA oligos. In CML cells, we observed inhibition of on- and off-target NHEJ. Depending on the inhibitor design used, indels could be virtually eliminated. Additionally, we used *in vitro* assays to show that a PAM-loop enables robust inhibition by decreasing the time required for inhibition to take place. This PAM-loop design was inspired by another study where DNA oligos were used to create a double-stranded PAM ⁵⁷. In contrast to most previously described Cas-inhibiting oligos ^{49,50}, the inhibitors described here inhibit in a guide sequence-dependent manner. Because of their sequence-specificity, these inhibitors theoretically allow modulating Cas9 activity at multiple sites independently. We found that the sequence-scrambled versions of the oligos (as well as the short (8 nt) complementary oligos) increased Cas9 activity, rather than inhibiting it. Perhaps the DNA oligos bind intracellular RNAs, preventing them from inhibiting RNP formation ⁵⁸.

We found that the 20nt oligo is the most potent inhibitor tested in CML cells. The 20nt oligo outperforms shorter inhibitors, most likely because it has higher affinity to the guide RNA. The PAM-loop is thought to increase affinity, partially rescuing inhibition by shorter oligos. Unexpectedly, the 8nt+PAM consistently showed a slightly different concentration response compared to the 20nt inhibitor (Fig. 2B, 3A). This might be partially explained by assuming that the 20nt inhibitor is cleaved by Cas9, while the 8nt+PAM is not. Indeed *in vitro* ssDNA cleavage by SpCas9 has been described ³ and DNA cleavage requires interactions between the protein and 'PAM distal' bases of the DNA ⁵⁹, which are not present in the 8nt or 8nt + PAM inhibitors. This might allow the 8nt+PAM inhibitor to evade cleavage and inhibit Cas9 near-optimally at an equimolar ratio to the RNP. In cases where inhibition should last, a shorter-than 20nt oligo (perhaps with a PAM-loop to improve affinity) might therefore be preferred. Alternatively, chemical modifications - already proven to yield potent Cas nuclease inhibitors ^{49,50} - can prevent cleavage of inhibitors.

In vitro, the 20nt inhibitor design did not show efficient inhibition when RNPs were pre-complexed. Indeed, the PAM-less inhibitors required more time to inhibit pre-formed RNP compared to the designs with a PAM-loop. *Ex vivo*, the Cas9 protein and guide RNA likely entered the cell as free protein and RNA. Therefore, additional time was required to form RNPs, possibly providing the 20nt inhibitor more time for strong inhibition. In addition, in the CML cells, the on-target DNA substrate is presumably at a relatively low concentration compared to the in the *in vitro* conditions and – once the RNP is formed - the time required for the Cas9 RNP to find its cognate target is much longer. This too would give more time to the 20nt design to establish strong inhibition, possibly explaining the observed differences between the *ex vivo* and *in vitro* experiments.

We found that the DNA oligo-based inhibitors affect both on- and off-targeting by Cas9 and can slightly increase specificity for activity on-target. However, the observed increase in specificity is minor and coincides with a substantial reduction of on-target activity. To put these results into context, we also tested engineered Cas9 variants (SpCas9-HF1⁶⁰ or Opti-SpCas9⁶¹) without inhibitors. We found that these variant Cas9 proteins are mostly more effective at increasing specificity than the inhibitors reported here (Supplementary Fig. 3). We thus conclude that for increasing specificity of Cas9, other measures (such as using engineered Cas9 protein variants) might be more suitable.

In an attempt to reveal the molecular basis of the specificity increase, we included conditions without inhibitors present where we varied the concentration of Cas9. The resulting data suggest that the observed increase in specificity could be a general consequence of lowering the concentration of active (cleavage-competent) RNPs, which has previously been described and explained by others^{15,62,63}. Unexpectedly, our highest Cas9 concentration led to increased on-target activity, but reduced off-target activity (Fig. 4A). The mechanism behind this is unknown and requires further research.

In CML cells, the most potent inhibition was observed with inhibitors that have the longest complementarity to the guide RNA. The *in vitro* experiments showed that a PAM-loop is required for rapid inhibition. Although the inhibitors performed slightly differently *in vitro* compared to *ex vivo*, in both settings DNA oligo-based inhibitors provided potent inhibition of Cas9 activity. Unmodified DNA oligos are inexpensive and easily manufactured and their design could easily be adapted to other RNA-guided nucleases. Overall, it is concluded that the guide-complementary DNA oligos reported here are promising candidates for inhibition of Cas9 activity.

materials and methods

Cell line and cell culture

The CML cell line used is a re-diploidized derivative of the HAP1 cell line, which was a kind gift of Dr. Thijn Brummelkamp⁶⁴. CML cells were cultured in Iscove's Modified Dulbecco's Medium (IMDM) (Gibco), supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin. Cells were grown at 37°C in a humidified atmosphere containing 5% CO₂.

Cas9 protein, guide RNAs, and inhibitors used *ex vivo*

Recombinant *S. pyogenes* Cas9 and high-fidelity protein variants were provided by Geijsen lab through Divvly (<https://divvly.com/geijsenlab>). guide RNAs used are synthetic guide RNAs which contain the target-specific crRNA and the scaffold tracrRNA (IDT). The crRNA and tracrRNA were dissolved in the Nuclease-free Duplex buffer (IDT) to reach the concentration of 200 µM. Equal volumes of dissolved crRNA and tracrRNA were mixed and annealed by heating for 5 minutes at 95°C and cooling down at room temperature. The oligo-based inhibitors used were unmodified single-stranded DNA oligos synthesized by IDT (IDT). Each of such oligos was dissolved in nuclease-free water to reach the concentration of 75 µM.

Induced transduction by osmocytosis and propanebetaine (iTOP)

The recombinant Cas9 proteins, guide RNAs, and oligo-based inhibitors were simultaneously transduced into CML cells by using the iTOP method as described previously⁵⁴. One day prior to the transduction, CML cells were plated at 18000 cells/well in the Matrigel-coated wells on 96-well plates, such that on the day of transduction, cells would reach about 70-80% confluence. Next day, for each well of the 96-well plate, 50 µL of iTOP mixture that contains 20 µL of transduction supplement (Opti-MEM media supplemented with 542 mM NaCl, 333 mM GABA, 1.67 x N2, 1.67 x B27, 1.67 x non-essential amino acids, 3.3 mM Glutamine, 167

ng/mL bFGF2, and 84 ng/mL EGF), 10 μ L of Cas9 protein (75 μ M), 7.5 μ L of guide RNA (100 μ M), 10 μ L of oligo-based inhibitors, and the excess volume of nuclease-free water to reach the 50- μ L total volume, were prepared. For the no-protein control, 10 μ L of protein storage buffer was used instead of the CRISPR nuclease protein; and for the no-guide control, the equal volume of nuclease-free water was used to replace the guide RNA or oligo-based inhibitors. The 50- μ L iTOP mixture was added onto the cells immediately after the culture medium was removed. The plate then was incubated in a cell culture incubator for 45 minutes, after which the iTOP mixture was gently removed and exchanged for 250 μ L of regular culture medium.

Isolation of genomic DNA in 96-well cell culture plates

Genomic DNA of each transduced cell samples was purified by direct in-plate cell lysis and DNA isolation according to a previously published protocol⁶⁵. Briefly, after aspirating culture media and adding 50 μ L of lysis buffer (10 mM Tris HCl pH 7.6, 10 mM EDTA, 100 mM NaCl, 0.5% N-Lauroylsarcosine sodium, 50 μ g/mL RNase A and 100 μ g/mL Protease K) to each well, the sealed plate was incubated at 55°C overnight; added 100 μ L of ice-cold NaCl-saturated ethanol (for 100 mL 100% ethanol, add 1.5 mL of 5M NaCl) to each well; allowed the plate stand still for 4 hours at room temperature to precipitate the DNA; washed the precipitated DNA with 75% ethanol for two times and let the plate air dry; DNA in each well was dissolved in 50 μ L TE buffer.

Deep-sequencing preparations

The regions of interest (supplementary table 1) were PCR amplified (supplementary table 3) from the extracted DNA using Q5 high-fidelity DNA polymerase (NEB). Of the primers used in the PCRs (supplementary table 2) the forward primers contained 5nt sequencing barcodes. All 576 amplifications were verified using 20g/L agarose gel electrophoresis. Equal volumes of the PCR products were pooled with samples from the same locus, but with unique barcodes. These 36 pools were then purified (Zymo Research Z4004) and quantified using Qubit dsDNA

BR (Thermo Fisher Scientific). Equal amounts of DNA from the samples were then pooled further to provide 6 pools where each sample contains a unique combination of sequencing barcode and amplified region. These samples were then sent to Baseclear B.V. for quality control, Index PCR, and sequencing on the NovaSeq 6000 for paired-end 150nt-long reads. The raw sequencing results are available on the NCBI Sequence Read Archive under BioProject PRJNA796802.

Analysis of deep-sequencing results

To analyze the deep-sequencing data the paired-end reads were programmatically merged using *seqprep*⁶⁶. Then the reads were filtered out that do not match the expected pattern of starting with a barcode followed by a forward primer annealing part, and ending with the associated reverse primer annealing part. We then split the reads by their barcodes and mapped the reads to the human genome (hg38) using *bowtie2*^{67,68}. The resulting alignments were then – using *samtools*^{69,70} - sorted, indexed, and split by the locus of interest where they align to the human genome. A separate script was used to score the amount of reads with insertions or deletions based on the CIGAR-strings from the alignment files. Another script was used to – for each indel – record its position and length. The used scripts were created in-house and will be made available upon reasonable request.

Calculation of specificity

Percentage specificity was calculated by dividing the *percentage reads with indels* of the on-target locus by the sum of *percentage reads with indels* across all three loci (on-target, off-target1, and off-target2). The resulting value answers the question ‘What percentage of the observed indels happen on-target?’, corrected for the total number of reads at each locus.

DNA cleavage assays

The relative inhibitory effect of the oligo-based inhibitors was determined using *S. pyogenes* Cas9 (NEB), Alt-R CRISPR-Cas9 sgRNA (IDT), EDTA (Merck) and proteinase K (NEB). The

substrate was obtained by PCR on CML cells derived gDNA using Q5 DNA polymerase (NEB) and purified on a 10g/L agarose gel using the Wizard SV Gel and PCR Clean-Up system (Promega). All mixtures were made in 1x NEBuffer 3.1 on ice with final concentrations of 80nM Cas9, 320nM sgRNA, 8 nM substrate and 1000nM oligo. For the pre-incubation assays Cas9 and sgRNA were mixed and preincubated for 20 minutes at 25°C and in the meantime the substrate and oligo mixes were made. Equal volumes of RNP and substrate+oligo mixes were combined and incubated at 37°C for 30 minutes. Reactions were stopped by placing them on ice and immediately adding 0.05V of 0.5M EDTA pH=8.0 and 0.05V of proteinase K. Time-point 0 samples were made by immediately stopping the reactions and storing them on ice during the incubation. After stopping samples were incubated at room temperature for 10 minutes and then TriTrack loading dye (ThermoFisher) was added to a final concentration of 1x and 12 µl was loaded on a 1% agarose gel. Gels were run for 45 minutes at 100V in 1xTAE buffer using SYBRSafe (ThermoFisher) as staining agent. Bands were visualized using the UVITEC Alliance (UVITEC) and quantified using the Image Lab software (Bio-Rad) version 6.0.1 build 34. Relative quantity of the remaining substrate was determined by setting the time-point 0 samples of each as 1 and measuring the relative intensity. % digestion was calculated as $(1-Irel)*100\%$. Assays were performed 3 times independently for each substrate.

For the non-preincubated in vitro assay the concentrations of all compounds were the same but the mixtures were substrate + Cas9 and oligo + sgRNA and they were combined on ice before incubation at 37°C.

Time-shift or time-lapse assay

For the time-lapse assay the same compounds and concentrations were used as for the in vitro assays but the preformed RNP and inhibitor were pre-incubated in 30 second steps before addition of the substrate. Therefore, three pre-mixes were made, Cas9 and sgRNA (RNP), oligo and substrate were diluted in 1x NEBuffer 3.1. The RNP complexes were preformed by incubation at 25°C for 20 minutes and 5 µl oligo dilution was put in 1.5 ml tubes (or

buffer in the case of the 'no inhibitor' control). All samples were put in a 37°C heat block and at different time-points 10 µl of the RNP or 5 µl substrate were added, so all samples had a digestion time of 30 minutes. Reactions were stopped and samples analyzed on gel as described above.

Electrophoresis mobility shift assay (EMSA)

Final concentrations of the compounds for EMSA were 170nM Cas9, 170nM sgRNA and 500nM oligo. Compounds were added together and pre-incubated at 37°C for 15 minutes after which buffer or the tested compound were added and the samples were incubated at 37°C for 15 minutes. Titrack loading dye was added and 5 µl was loaded on a 5% polyacrylamide gel in 0.5xTB buffer (45 mM Tris, 45 mM boric acid) and ran at 15mA for 40 minutes. The gel was stained with SYBRGold (ThermoFischer) for 5 minutes, destained in 0.5xTB buffer for 10 minutes and visualized using the UVITEC Alliance (UVITEC).

Data visualization and statistics

The data presented in this study were visualized using the *datavis.ipynb* Jupyter Notebook file. In addition to *Jupyter Notebook*⁷², we used *Python3*⁷¹, *numpy*⁷³, *pandas*⁷⁴, *matplotlib*⁷⁵, and *seaborn*⁷⁶. In cases where results were quantified, triplicates were used for each condition. Where statistics were indicated, those were done by conducting a Shapiro-Wilk test for normality, a Levene's test for equal variance, a one-way ANOVA, and lastly a Tukey's range posthoc test. For these statistics, we used *SciPy*⁷⁷ and *statsmodels*⁷⁸ in addition to the programs and packages used for the data visualization.

Literature

1. Barrangou, R. *et al.* CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science* **315**, 1709–1712 (2007).
2. van der Oost, J., Jore, M. M., Westra, E. R., Lundgren, M. & Brouns, S. J. J. CRISPR-based adaptive and heritable immunity in prokaryotes. *Trends Biochem. Sci.* **34**, 401–407 (2009).
3. Jinek, M. *et al.* A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **337**, 816–821 (2012).
4. Zetsche, B. *et al.* Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell* **163**, 759–771 (2015).
5. Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67 (2014).
6. Yamano, T. *et al.* Crystal Structure of Cpf1 in Complex with Guide RNA and Target DNA. *Cell* **165**, 949–962 (2016).
7. Gao, L. *et al.* Engineered Cpf1 variants with altered PAM specificities. *Nat. Biotechnol.* **35**, 789–792 (2017).
8. Swarts, D. C., van der Oost, J. & Jinek, M. Structural Basis for Guide RNA Processing and Seed-Dependent DNA Targeting by CRISPR-Cas12a. *Mol. Cell* **66**, 221–233.e4 (2017).
9. Jeon, Y. *et al.* Direct observation of DNA target searching and cleavage by CRISPR-Cas12a. *Nat. Commun.* **9**, 1–11 (2018).
10. Jiang, W., Bikard, D., Cox, D., Zhang, F. & Marraffini, L. A. RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nat. Biotechnol.* **31**, 233–239 (2013).
11. Cong, L. *et al.* Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **339**, 819–823 (2013).
12. Scully, R., Panday, A., Elango, R. & Willis, N. A. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat. Rev. Mol. Cell Biol.* **20**, 698–714 (2019).
13. Ceccaldi, R., Rondinelli, B. & D’Andrea, A. D. Repair Pathway Choices and Consequences at the Double-Strand Break. *Trends Cell Biol.* **26**, 52–64 (2016).
14. Mali, P. *et al.* CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nat. Biotechnol.* **31**, 833–838 (2013).

15. Hsu, P. D. *et al.* DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat. Biotechnol.* **31**, 827–832 (2013).
16. Pattanayak, V. *et al.* High-throughput profiling of off-target DNA cleavage reveals RNA-programmed Cas9 nuclease specificity. *Nat. Biotechnol.* **31**, 839–843 (2013).
17. Fu, Y. *et al.* High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* **31**, 822–826 (2013).
18. Cho, S. W. *et al.* Analysis of off-target effects of CRISPR/Cas-derived RNA-guided endonucleases and nickases. *Genome Res.* **24**, 132–141 (2014).
19. Bondy-Denomy, J., Pawluk, A., Maxwell, K. L. & Davidson, A. R. Bacteriophage genes that inactivate the CRISPR/Cas bacterial immune system. *Nature* **493**, 429–432 (2013).
20. Bondy-Denomy, J. *et al.* Multiple mechanisms for CRISPR–Cas inhibition by anti-CRISPR proteins. *Nature* **526**, 136–139 (2015).
21. Marino, N. D., Pinilla-Redondo, R., Csörgő, B. & Bondy-Denomy, J. Anti-CRISPR protein applications: natural brakes for CRISPR-Cas technologies. *Nat. Methods* 1–9 (2020) doi:10.1038/s41592-020-0771-6.
22. Pawluk, A. *et al.* Naturally Occurring Off-Switches for CRISPR-Cas9. *Cell* **167**, 1829-1838.e9 (2016).
23. Rauch, B. J. *et al.* Inhibition of CRISPR-Cas9 with Bacteriophage Proteins. *Cell* **168**, 150-158.e10 (2017).
24. Shin, J. *et al.* Disabling Cas9 by an anti-CRISPR DNA mimic. *Sci. Adv.* **3**, (2017).
25. Hynes, A. P. *et al.* Widespread anti-CRISPR proteins in virulent bacteriophages inhibit a range of Cas9 proteins. *Nat. Commun.* **9**, 2919 (2018).
26. Trasanidou, D. *et al.* Keeping crispr in check: diverse mechanisms of phage-encoded anti-crisprs. *FEMS Microbiol. Lett.* **366**, (2019).
27. Wu, X., Huang, H., Yu, B. & Zhang, J. A Blue Light-Inducible CRISPR-Cas9 System for Inhibiting Progression of Melanoma Cells. *Front. Mol. Biosci.* **7**, 366 (2020).
28. Polstein, L. R. & Gersbach, C. A. A light-inducible CRISPR/Cas9 system for control of endogenous gene activation. *Nat. Chem. Biol.* **11**, 198–200 (2015).

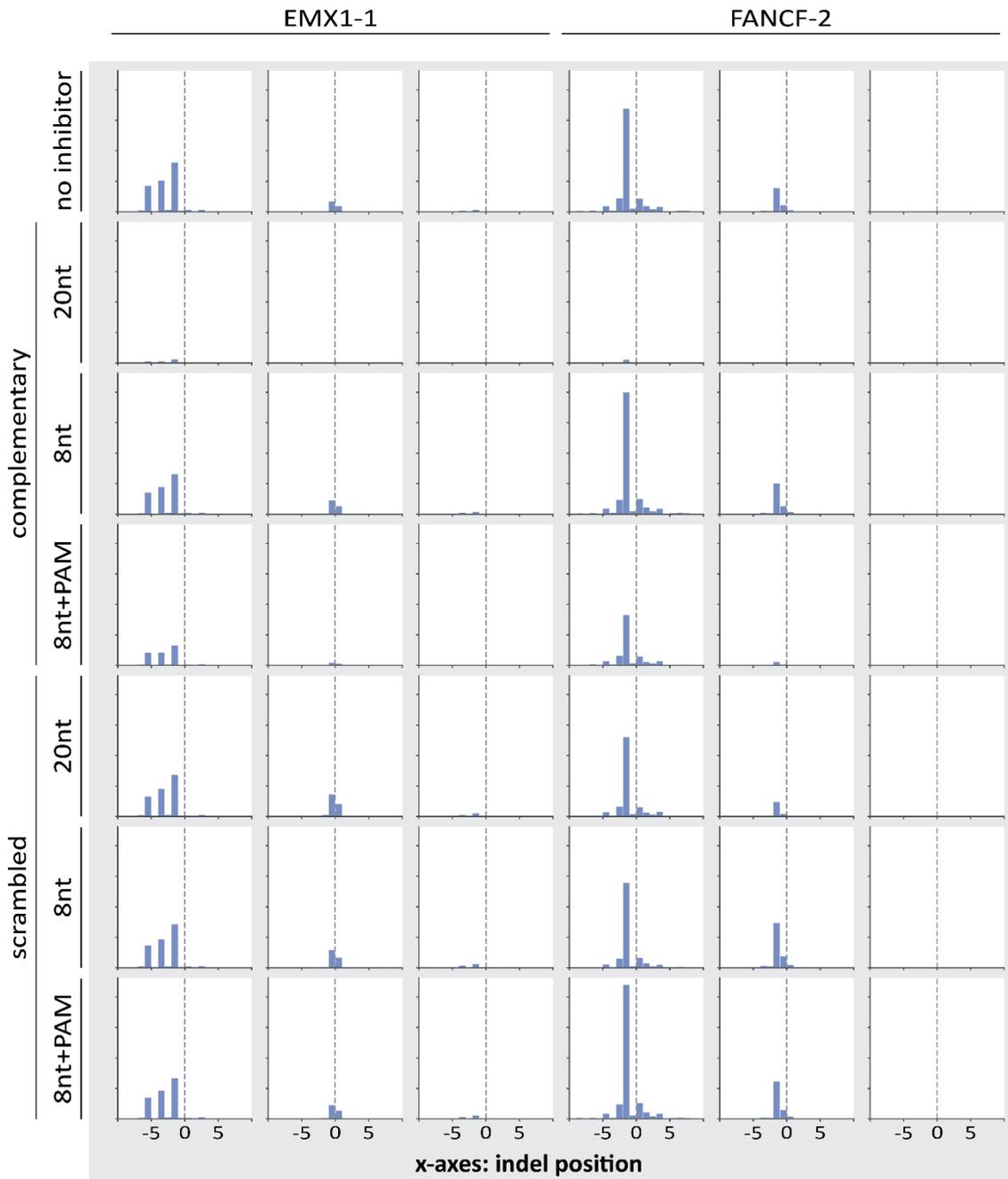
29. DeLorenzo, D. M., Rottinghaus, A. G., Henson, W. R. & Moon, T. S. Molecular Toolkit for Gene Expression Control and Genome Modification in *Rhodococcus opacus* PD630. *ACS Synth. Biol.* **7**, 727–738 (2018).
30. Liu, Y. *et al.* Directing cellular information flow via CRISPR signal conductors. *Nat. Methods* **13**, 938–944 (2016).
31. Koopal, B., Kruis, A. J., Claassens, N. J., Nobrega, F. L. & van der Oost, J. Incorporation of a Synthetic Amino Acid into dCas9 Improves Control of Gene Silencing. *ACS Synth. Biol.* **8**, 216–222 (2019).
32. Patinios, C. *et al.* *SIBR-Cas enables host-independent and universal CRISPR genome engineering in bacteria.* 2021.04.26.441145
<https://www.biorxiv.org/content/10.1101/2021.04.26.441145v1> (2021)
doi:10.1101/2021.04.26.441145.
33. Nishimasu, H. *et al.* Crystal Structure of *Staphylococcus aureus* Cas9. *Cell* **162**, 1113–1126 (2015).
34. Chew, W. L. *et al.* A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat. Methods* **13**, 868–874 (2016).
35. Zetsche, B., Volz, S. E. & Zhang, F. A split-Cas9 architecture for inducible genome editing and transcription modulation. *Nat. Biotechnol.* **33**, 139–142 (2015).
36. Gao, Y. *et al.* Complex transcriptional modulation with orthogonal and inducible dCas9 regulators. *Nat. Methods* **13**, 1043–1049 (2016).
37. Moreno-Mateos, M. A. *et al.* CRISPR-Cpf1 mediates efficient homology-directed repair and temperature-controlled genome editing. *Nat. Commun.* **8**, 1–9 (2017).
38. Truong, D.-J. J. *et al.* Development of an intein-mediated split-Cas9 system for gene therapy. *Nucleic Acids Res.* **43**, 6450–6458 (2015).
39. Yu, Y. *et al.* Engineering a far-red light-activated split-Cas9 system for remote-controlled genome editing of internal organs and tumors. *Sci. Adv.* **6**, eabb1777 (2020).
40. Richter, F. *et al.* Engineering of temperature- and light-switchable Cas9 variants. *Nucleic Acids Res.* **44**, 10003–10014 (2016).
41. Konermann, S. *et al.* Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**, 583–588 (2015).

42. Nguyen, D. P. *et al.* Ligand-binding domains of nuclear receptors facilitate tight control of split CRISPR activity. *Nat. Commun.* **7**, 12009 (2016).
43. Hemphill, J., Borchardt, E. K., Brown, K., Asokan, A. & Deiters, A. Optical Control of CRISPR/Cas9 Gene Editing. *J. Am. Chem. Soc.* **137**, 5642–5645 (2015).
44. Nihongaki, Y., Kawano, F., Nakajima, T. & Sato, M. Photoactivatable CRISPR-Cas9 for optogenetic genome editing. *Nat. Biotechnol.* **33**, 755–760 (2015).
45. Oakes, B. L. *et al.* Profiling of engineering hotspots identifies an allosteric CRISPR-Cas9 switch. *Nat. Biotechnol.* **34**, 646–651 (2016).
46. Wright, A. V. *et al.* Rational design of a split-Cas9 enzyme complex. *Proc. Natl. Acad. Sci.* **112**, 2984–2989 (2015).
47. Davis, K. M., Pattanayak, V., Thompson, D. B., Zuris, J. A. & Liu, D. R. Small molecule-triggered Cas9 protein with improved genome-editing specificity. *Nat. Chem. Biol.* **11**, 316–318 (2015).
48. Chow, R. D., Kim, H. R. & Chen, S. Programmable sequential mutagenesis by inducible Cpf1 crRNA array inversion. *Nat. Commun.* **9**, 1903 (2018).
49. Li, B. *et al.* Synthetic Oligonucleotides Inhibit CRISPR-Cpf1-Mediated Genome Editing. *Cell Rep.* **25**, 3262-3272.e3 (2018).
50. Barkau, C. L., O'Reilly, D., Rohilla, K. J., Damha, M. J. & Gagnon, K. T. Rationally Designed Anti-CRISPR Nucleic Acid Inhibitors of CRISPR-Cas9. *Nucleic Acid Ther.* **29**, 136–147 (2019).
51. Rose, J. C. *et al.* Suppression of unwanted CRISPR-Cas9 editing by co-administration of catalytically inactivating truncated guide RNAs. *Nat. Commun.* **11**, 2697 (2020).
52. Kiani, S. *et al.* Cas9 gRNA engineering for genome editing, activation and repression. *Nat. Methods* **12**, 1051–1054 (2015).
53. Dahlman, J. E. *et al.* Orthogonal gene knockout and activation with a catalytically active Cas9 nuclease. *Nat. Biotechnol.* **33**, 1159–1161 (2015).
54. D'Astolfo, D. S. *et al.* Efficient Intracellular Delivery of Native Proteins. *Cell* **161**, 674–690 (2015).
55. Slaymaker, I. M. *et al.* Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88 (2016).
56. Chen, J. S. *et al.* Enhanced proofreading governs CRISPR–Cas9 targeting accuracy. *Nature* **550**, 407–410 (2017).

57. O'Connell, M. R. *et al.* Programmable RNA recognition and cleavage by CRISPR/Cas9. *Nature* **516**, 263–266 (2014).
58. Kagita, A. *et al.* Efficient ssODN-Mediated Targeting by Avoiding Cellular Inhibitory RNAs through Precomplexed CRISPR-Cas9/sgRNA Ribonucleoprotein. *Stem Cell Rep.* **16**, 985–996 (2021).
59. Yang, M. *et al.* The Conformational Dynamics of Cas9 Governing DNA Cleavage Are Revealed by Single-Molecule FRET. *Cell Rep.* **22**, 372–382 (2018).
60. Kleinstiver, B. P. *et al.* High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature* **529**, 490–495 (2016).
61. Choi, G. C. G. *et al.* Combinatorial mutagenesis en masse optimizes the genome editing activities of SpCas9. *Nat. Methods* **16**, 722–730 (2019).
62. Bisaria, N., Jarmoskaite, I. & Herschlag, D. Lessons from Enzyme Kinetics Reveal Specificity Principles for RNA-Guided Nucleases in RNA Interference and CRISPR-Based Genome Editing. *Cell Syst.* **4**, 21–29 (2017).
63. Herschlag, D. Implications of ribozyme kinetics for targeting the cleavage of specific RNA molecules in vivo: more isn't always better. *Proc. Natl. Acad. Sci.* **88**, 6921–6925 (1991).
64. Kotecki, M., Reddy, P. S. & Cochran, B. H. Isolation and characterization of a near-haploid human cell line. *Exp. Cell Res.* **252**, 273–280 (1999).
65. Ramírez-Solis, R. *et al.* Genomic DNA microextraction: a method to screen numerous samples. *Anal. Biochem.* **201**, 331–335 (1992).
66. St. John, J. *SeqPrep.* (2016).
67. Langmead, B. *bowtie2.* (2019).
68. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
69. *SAMtools.* (samtools, 2021).
70. Li, H. *et al.* The Sequence Alignment/Map format and SAMtools. *Bioinforma. Oxf. Engl.* **25**, 2078–2079 (2009).
71. Van Rossum, G. & Drake, F. L. *Python 3 Reference Manual.* (CreateSpace, 2009).
72. Kluyver, T. *et al.* Jupyter Notebooks – a publishing format for reproducible computational workflows. in *Positioning and Power in Academic Publishing: Players, Agents and Agendas* (eds. Loizides, F. & Schmidt, B.) 87–90 (IOS Press, 2016).

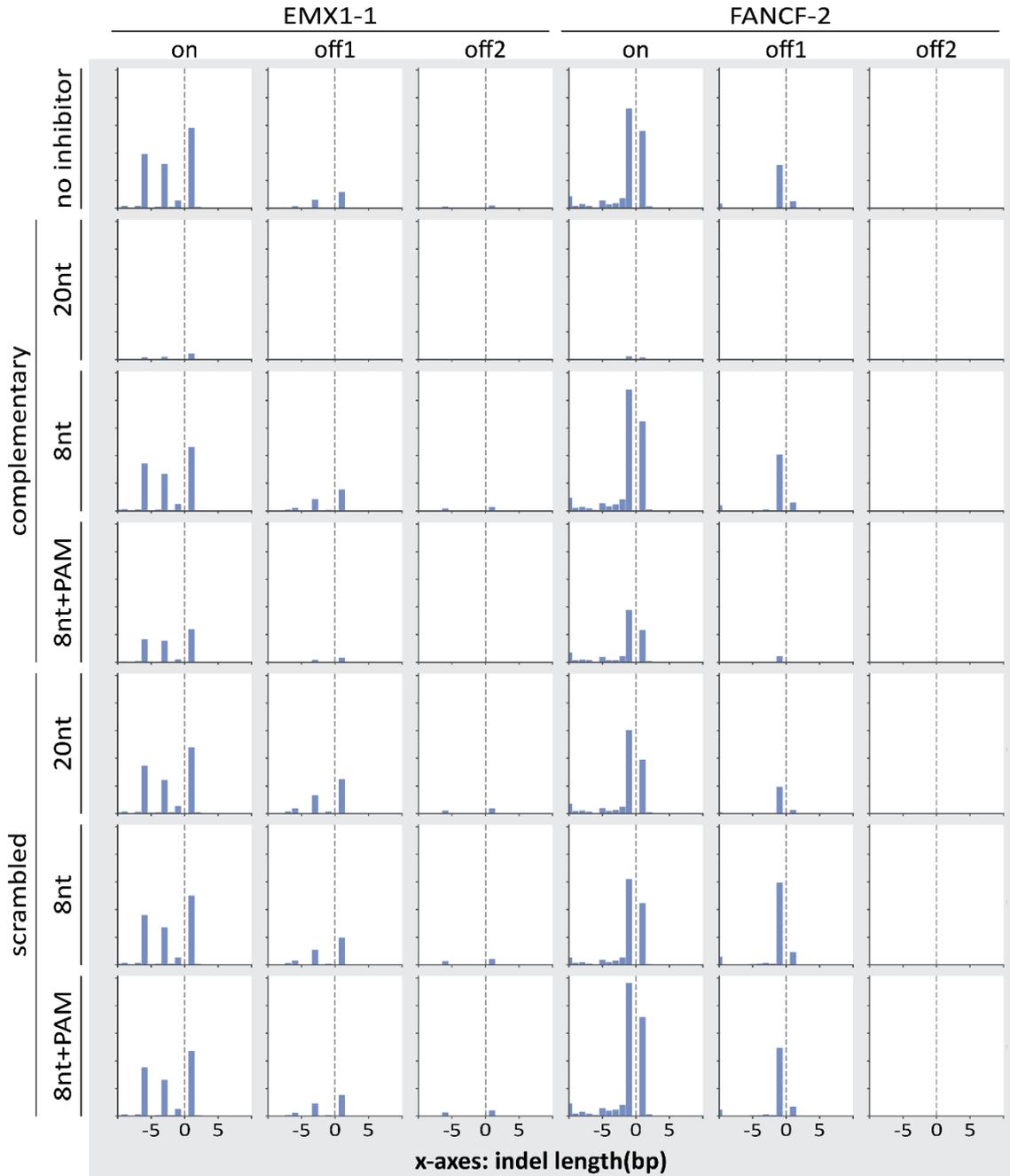
73. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362 (2020).
74. *pandas*. (Zenodo, 2020). doi:10.5281/zenodo.3509134.
75. Hunter, J. D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).
76. Waskom, M. L. seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 3021 (2021).
77. Virtanen, P. *et al.* SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **17**, 261–272 (2020).
78. Seabold, S. & Perktold, J. statsmodels: Econometric and statistical modeling with python. in *9th Python in Science Conference* (2010).

Supplementary information



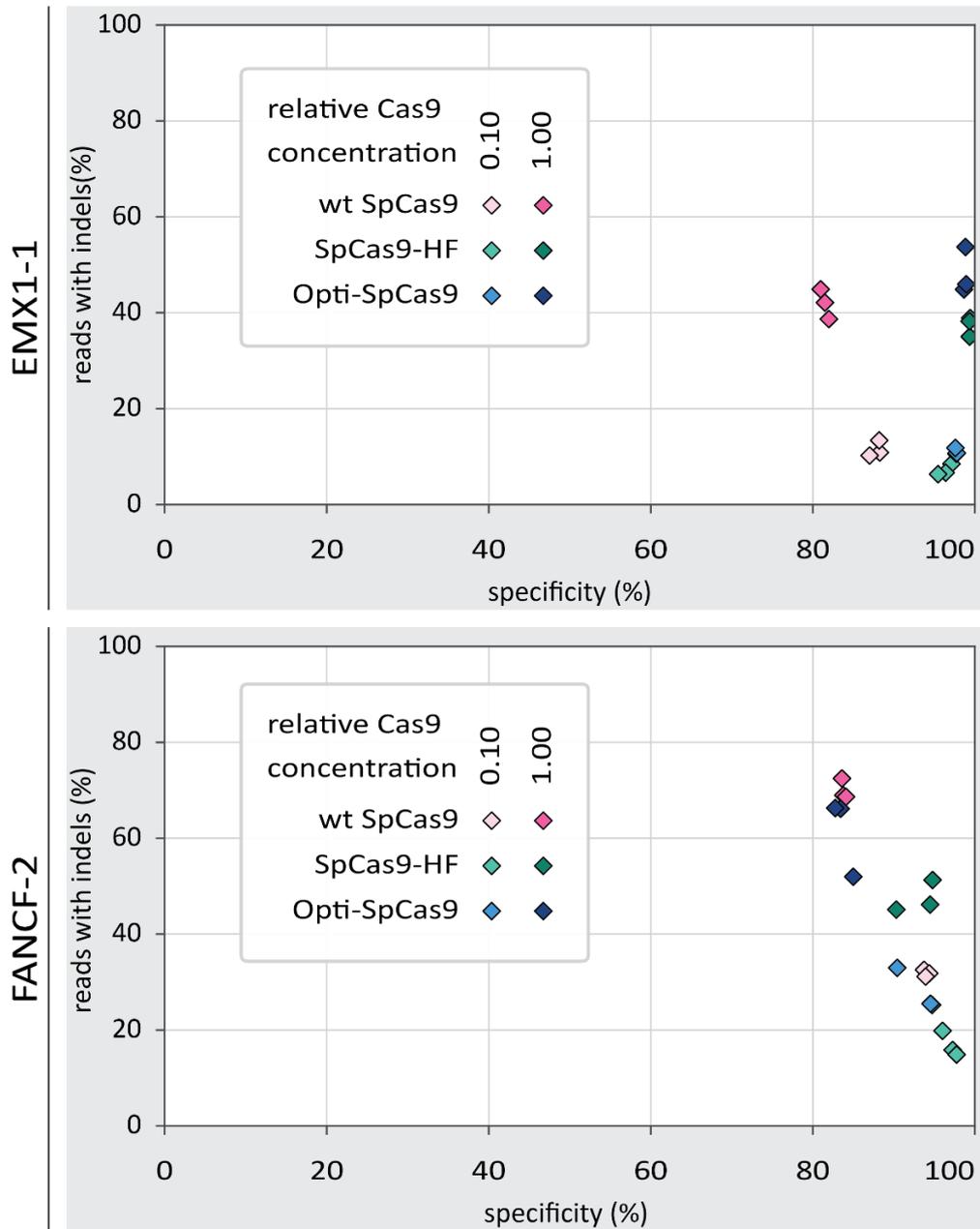
Supplementary Fig. 1. Indel positions distribution.

Histograms of the positions of the indels for different DNA oligo-based designs. The '0' position is the Cas9 cut-site for each amplicon. We only displayed indels that occurred within 9 bp from the cut-site. The DNA oligos included in these graphs were delivered at molar concentrations equal to the concentration of Cas9 and guide RNA.



Supplementary Fig. 2. Indel length distribution.

Histograms of the lengths of the indels for different DNA oligo-based designs. Negative lengths indicate deletions, whereas positive lengths indicate insertions. A length of '0' would not be an indel, and was not included in these graphs. We only displayed insertions or deletions of 10 bp or shorter. The DNA oligos included in these graphs were delivered at molar concentrations equal to the concentration of Cas9 and guide RNA.



Supplementary Fig. 3. Specificity and activity of engineered Cas9 variants.

comparison of *percentage reads with indels* and *percentage specificity* for different Cas9 variants. For each protein, and for each protein concentration, individual replicates are displayed as diamonds with the same color.

Supplementary Table 1. Genomic loci.

The genomic loci that were targeted in the *ex vivo* experiment, and which were amplified to provide DNA substrate for the *in vitro* experiments. For each target we list whether it is included as an on- or off-target site, which on-target site it relates to, the nucleotide sequence, the PAM, the chromosome on which it is located, the coordinate of where on the chromosome it is located, and finally references to sources, based on which we decided to include these loci.

on/off	on-target	sequence	PAM	chr	coordinate	refs
on	EMX1-1	GAGTCCGAGCAGAAGAAGAA	gGG	2	72933853	55
off1	EMX1-1	GAGTTAGAGCAGAAGAAGAA	aGG	5	45358962	55
off2	EMX1-1	GAGTCTAAGCAGAAGAAGAA	gAG	15	43817549	55
on	FANCF-2	GCTGCAGAAGGGATTCCATG	aGG	11	22625792	56
off1	FANCF-2	GCTGCAGAAGGGATTCCAAG	gGG	22	36556948	56
off2	FANCF-2	GACGCAGAAGGGACTCCATG	gGG	6	27764843	56

Supplementary Table 2. Deep-sequencing preparation primers.

The primers that were used to create the amplicons that were sent for deep-sequencing. In the sequences, '---' is used to indicate the upstream sequences that Baseclear B.V. requested for further preparation of the sequencing samples. These sequences are excluded from the amplicon length in the last column. 'NNNNN' is used to indicate the 5 nt long barcodes that we used to distinguish between conditions.

primer	sequence	annealing temperature
EMX1-1_on_fw	---NNNNNAAGGTGTGGTTCAGAACC	66°C
EMX1-1_on_rv	---CGATGTCTCCCATTTGG	
EMX1-1_off-1_fw	---NNNNNTTACCATAGACTATCACC	54°C
EMX1-1_off-1_rv	---GGTTACAGAAAGAATAGG	
EMX1-1_off-2_fw	---NNNNNTGTGCTTCAACCCATCACG	67°C
EMX1-1_off-2_rv	---GCAGTCTCATGACTTGGCC	
FANCF-2_on_fw	---NNNNNCGTAGGTAGTGCTTGAGACC	65°C
FANCF-2_on_rv	---CATTTTCGCGGATGTTCCAATC	
FANCF-2_off-1_fw	---NNNNNCCATTTCTGTCTCCACCTCC	62°C
FANCF-2_off-1_rv	---CCTCTCTCTCCACCGAGTTAC	
FANCF-2_off-2_fw	---NNNNNGTTAATGTACAAGGGGTGAG	61°C
FANCF-2_off-2_rv	---CAATCCAGGGCCCTATCTC	

Supplementary table 3. Deep-sequencing preparation PCR conditions.

Left: The components, their concentrations and used volumes for the 25µL PCR reactions.

Right: The thermocycler program used for the PCRs.

component	volume (ul)	temp. (°C)	duration	
2x Q5 mastermix	12.5	98	30s	25x
fw primer 2.5µM	5.0	98	10s	
rv primer 25µM	0.5	(supp. table 2)	30s	
DNA 28ng/µL	7	72	15s	
total	25	72	2min	

Supplementary table 4. *In vitro* DNA substrate primers

The primers used to amplify the 1500bp DNA substrates from the human genome for the *in vitro* assays. Cleavage by Cas9 results in two fragments of lengths 500bp and 1000bp.

Locus	fw primer	rv primer
EMX1-1	tctcatttactactcacatccactctg	agggcttaaggctgagcc
FANCF-2	cgctttacaggtctccagg	cctattaatgccaggcgctatg

Supplementary table 5. DNA oligo-based inhibitors

The names, sequences, and lengths of the DNA oligonucleotides that were used as inhibitors. In the sequences, both the guide-complementary part and the PAM are in uppercase, but the loop connecting them is lowercase.

site	inhibitor name	DNA oligo sequence	length (nt)
EMX1-1	8nt + PAM	AGGcaaagcctTTCTTCTT	20
	20nt	TTCTTCTTCTGCTCGACTC	20
	off 1 20nt	TTCTTCTTCTGCTCTAACTC	20
	off 2 20nt	TTCTTCTTCTGCTTAGACTC	20
	8nt	TTCTTCTT	8
	scrambled 8nt + PAM	AGGcaaagcctCTTTCTTT	20
	scrambled 20nt	CTGTACGTGCCCTCTTTCTT	20
	scrambled 8nt	CTTTCTTT	8
FANCF-2	8nt + PAM	AGGcaaagcctCATGGAAT	20
	20nt	CATGGAATCCCTTCTGCAGC	20
	off 1 20nt	CTTGGAATCCCTTCTGCAGC	20
	off 2 20nt	CATGGAGTCCCTTCTGCGTC	20
	8nt	CATGGAAT	8
	scrambled 8nt + PAM	AGGcaaagcctAGTCAAGT	20
	scrambled 20nt	AGAATCCCACGCTTCCGTTG	20
	scrambled 8nt	AGTCAAGT	8