



# MIT Open Access Articles

## *Structure-guided chemical modification of guide RNA enables potent non-viral in vivo genome editing*

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

<b>Citation</b>	Yin, Hao et al. "Structure-guided chemical modification of guide RNA enables potent non-viral in vivo genome editing." Nature Biotechnology 35, 12 (November 2017): 1179-1187 © 2017 Nature Publishing Group
<b>As Published</b>	<a href="http://dx.doi.org/10.1038/nbt.4005">http://dx.doi.org/10.1038/nbt.4005</a>
<b>Publisher</b>	Springer Science and Business Media LLC
<b>Version</b>	Author's final manuscript
<b>Citable link</b>	<a href="https://hdl.handle.net/1721.1/121981">https://hdl.handle.net/1721.1/121981</a>
<b>Terms of Use</b>	Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.



# HHS Public Access

Author manuscript

*Nat Biotechnol.* Author manuscript; available in PMC 2018 April 16.

Published in final edited form as:

*Nat Biotechnol.* 2017 December ; 35(12): 1179–1187. doi:10.1038/nbt.4005.

## Structure-guided chemical modification of guide RNA enables potent non-viral *in vivo* genome editing

Hao Yin<sup>1</sup>, Chun-Qing Song<sup>2,12</sup>, Sneha Suresh<sup>1,12</sup>, Qiongqiong Wu<sup>1,12</sup>, Stephen Walsh<sup>1</sup>, Luke Hyunsik Rhym<sup>1,3</sup>, Esther Mintzer<sup>4</sup>, Mehmet Fatih Bolukbasi<sup>4,5</sup>, Lihua Julie Zhu<sup>4,6,7</sup>, Kevin Kauffman<sup>1,3</sup>, Haiwei Mou<sup>2</sup>, Alicia Oberholzer<sup>1</sup>, Junmei Ding<sup>1</sup>, Suet-Yan Kwan<sup>2</sup>, Roman L Bogorad<sup>1</sup>, Timofei Zatsepin<sup>8,9</sup>, Victor Koteliansky<sup>8</sup>, Scot A Wolfe<sup>4,5</sup>, Wen Xue<sup>2,4,7</sup>, Robert Langer<sup>1,3,10,11</sup>, and Daniel G Anderson<sup>1,3,10,11</sup>

<sup>1</sup>David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

<sup>2</sup>RNA Therapeutics Institute, University of Massachusetts Medical School, Worcester, Massachusetts, USA

<sup>3</sup>Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

<sup>4</sup>Department of Molecular, Cell and Cancer Biology, University of Massachusetts Medical School, Worcester, Massachusetts, USA

<sup>5</sup>Department of Biochemistry and Molecular Pharmacology, University of Massachusetts Medical School, Worcester, Massachusetts, USA

<sup>6</sup>Program in Bioinformatics and Integrative Biology, University of Massachusetts Medical School, Worcester, Massachusetts, USA

<sup>7</sup>Department of Molecular Medicine, University of Massachusetts Medical School, Worcester, Massachusetts, USA

<sup>8</sup>Center of Translational Biomedicine, Skolkovo Institute of Science and Technology, Skolkovo, Moscow, Russia

<sup>9</sup>Department of Chemistry, Lomonosov Moscow State University, Moscow, Russia

<sup>10</sup>Harvard-MIT Division of Health Sciences & Technology, Cambridge, Massachusetts, USA

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

Correspondence should be addressed to D.G.A. (dgander@mit.edu).

<sup>12</sup>These authors contributed equally to this work.

Note: Any Supplementary Information and Source Data files are available in the online version of the paper.

### AUTHOR CONTRIBUTIONS

H.Y. conceived of and designed the study, and directed the project. H.Y., S.S., C.-Q.S., S.W., L.H.R., K.K., H.M., Q.W., E.M., M.F.B., L.J.Z., S.-Y.K., T.Z., S.A.W., A.O., J.D., R.L.B. and W.X. performed experiments and analyzed data. C.-Q.S. made the figures with H.Y. W.X., V.K. and R.L. provided conceptual advice. H.Y. wrote the manuscript with comments from all authors. D.G.A. supervised the project.

### COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

<sup>11</sup>Institute of Medical Engineering and Science, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

## Abstract

Efficient genome editing with Cas9–sgRNA *in vivo* has required the use of viral delivery systems, which have limitations for clinical applications. Translational efforts to develop other RNA therapeutics have shown that judicious chemical modification of RNAs can improve therapeutic efficacy by reducing susceptibility to nuclease degradation. Guided by the structure of the Cas9–sgRNA complex, we identify regions of sgRNA that can be modified while maintaining or enhancing genome-editing activity, and we develop an optimal set of chemical modifications for *in vivo* applications. Using lipid nanoparticle formulations of these enhanced sgRNAs (e-sgRNA) and mRNA encoding Cas9, we show that a single intravenous injection into mice induces >80% editing of *Pcsk9* in the liver. Serum *Pcsk9* is reduced to undetectable levels, and cholesterol levels are significantly lowered about 35% to 40% in animals. This strategy may enable non-viral, Cas9-based genome editing in the liver in clinical settings.

---

CRISPR-based genome editing has considerable therapeutic potential in a wide range of diseases<sup>1–5</sup>. A key challenge is achieving efficient, clinically suitable delivery of genome editing macromolecules. CRISPR–Cas9, a complex of the RNA-guided nuclease (Cas9) and a single guide RNA (sgRNA)<sup>2</sup>, recognizes a protospacer-adjacent motif (PAM) and a 20-nucleotide sequence in the genome by Watson–Crick base pairing<sup>2–5</sup> and generates double-stranded DNA breaks (DSBs), which are repaired by endogenous cellular mechanisms such as homology-directed repair (HDR) and nonhomologous end-joining (NHEJ)<sup>5</sup>. Long-term expression of Cas9–sgRNA *in vivo* has been shown using viral vectors, including adeno-associated virus (AAV)<sup>6–9</sup>. However, an ideal CRISPR–Cas9 delivery system would limit the duration of exposure to the editing machinery to minimize potential off-target effects<sup>10</sup>. In addition, spCas9, the most commonly used form of Cas9, is difficult to fit in typical AAV constructs with strong promoters<sup>6</sup>. Although a smaller form of Cas9 was packed into a single AAV construct<sup>11</sup>, concerns regarding potential off-target effects remain when Cas9 is stably expressed by AAV delivery<sup>12</sup>. Moreover, the immune response to AAV capsids can limit repeat dosing in patients<sup>13,14</sup>, and the long-term presence of Cas9, a bacterial protein, in human tissue also increases the risk of immunogenicity<sup>13,15</sup>. These limitations can be addressed using non-viral delivery systems<sup>12</sup>. The use of Cas9–sgRNA ribonucleoprotein (RNP) has been tested in cell culture and for local delivery in mouse inner ear cells<sup>16</sup> but has not yet been demonstrated for systemic *in vivo* delivery. Previously, we used lipid nanoparticle (LNP)-encapsulated Cas9 mRNA in combination with an AAV carrying a sgRNA and a repair template for efficient genome editing in the mouse liver<sup>12</sup>. However, to our knowledge, a fully non-viral, systemic Cas9 genome editing system allowing efficient *in vivo* gene modification has not been described<sup>17</sup>.

Efforts to translate RNA therapeutics, in particular short interfering RNA (siRNA) and anti-sense oligonucleotides (ASO), have led to the identification of a number of chemical modifications to RNA that substantially improve efficacy of gene knockdown following *in vivo* delivery<sup>18,19</sup>. RNAs are particularly sensitive to nucleases in the blood and in cells following administration<sup>20</sup>. Two examples of sugar modifications are the 2′ *O*-methyl

ribonucleotide (2'OMe RNA) and 2'-deoxy-2'-fluoro-ribonucleotide (2'F RNA) forms, which reduce susceptibility to nucleases<sup>18</sup>. In addition, a chemical modification of the phosphate backbone, by inclusion of thioester groups, substantially increases nuclease resistance and thereby improves *in vivo* function<sup>18</sup>. However, chemical modifications can also interfere with the biological function of therapeutic RNAs<sup>21</sup>. For example, certain chemical modifications can interfere with the activity of the siRNA-Argonaut complex by preventing RNA-induced silencing complexes (RISCs) from stably interacting with siRNA<sup>21</sup>. Similarly, ASOs, which function through the use of RNase H, show a different pattern of sensitivity to chemical modifications<sup>19</sup>.

Targeting of the Cas9 complex is guided by sgRNAs, which combines CRISPR RNA (crRNA) and *trans*-activating crRNA (tracrRNA)<sup>2</sup>. These sgRNAs are about 100 nucleotides (nt) long, with 20 nt at the 5' end hybridizing the complementary DNA sequence and the remaining structure for Cas9 recognition and binding<sup>2-5</sup>. Chemical modification of three nucleotides at both the 5' and 3' ends has been shown to improve genome editing efficiencies *in vitro* in human cells<sup>22</sup>. Rational chemical modifications of crRNA in combination with unmodified tracrRNA mediated effective genome editing in human cells<sup>23</sup>. However, in these studies the majority of the RNA remained unmodified, and the potential of using these chemical modifications *in vivo* was not explored<sup>22,23</sup>. We reasoned that *in vivo* application of sgRNA would benefit from focused chemical modification.

Guided by the structure of the Cas9–sgRNA complex, we here develop enhanced sgRNA (e-sgRNA), defined as 70 out of 101 nt of sgRNA modified with a 2' hydroxyl (OH) group, and a number of phosphorothioate bonds. We identify regions of sgRNA that tolerate chemical modification without inhibiting the interaction of Cas9 and sgRNA while maintaining or enhancing genome editing activity. Using these modifications, we demonstrate that a single dose of the formulated e-sgRNA and Cas9 mRNA allows for nearly complete editing of the target gene in hepatocytes *in vivo*.

## RESULTS

### Structure-guided chemical modification of sgRNA's invariable part

To study the modifications tolerated by sgRNA, we separately analyzed its invariable and variable parts. The variable part, as the guide sequences, is the region of 1–20 nucleotides from the 5' end (Fig. 1a). The invariable part is 21 to 101 nucleotides from the 5' end, including the bona fide Cas9 protein-binding region and the tail region (Fig. 1a). We first synthesized sgRNAs containing the guide sequence targeting GFP protein<sup>24</sup>.

To more rapidly identify modifications that were tolerated by the sgRNA, we made use of a cell reporter system to test the editing efficiency of modified sgRNAs. HEK293 cells were engineered to stably express GFP and spCas9 (Fig. 1a). Introducing a functional sgRNA targeting GFP abrogates the expression of GFP through the generation of frameshifting indel mutations (Fig. 1a). If a pattern of chemical modifications on the sgRNA is not well tolerated, such chemically modified sgRNA induces fewer GFP-negative cells (GFP<sup>-</sup>) than unmodified sequence. We generated a number of sgRNAs modified with 2'F RNA, 2'OMe RNA and the phosphorothioate bond (PS)<sup>25</sup> (Fig. 1b) at different positions (Fig. 1c). Native

sgRNA targeting GFP efficiently generated GFP<sup>-</sup> cells (82.3% ± 1.7%) with a relatively high dose of sgRNA transfected (1 µg/mL; Fig. 1c and Supplementary Fig. 1). By contrast, 2'F modification of “U” and “C” on the Cas9 binding region nearly abolished the editing effect of the GFP sgRNA. Furthermore, when we substituted all of the U and C with 2'F modified U and C in the Cas9 binding region and tail region (2'F, U&C mod 21-101nt), the editing ability of the sgRNA was totally eliminated, as indicated by background level of GFP<sup>-</sup> cells after transfection (Fig. 1c). To examine whether the Cas9 binding and tail region can tolerate full substitutions of other chemical modifications, we introduced 2'OMe (2'OMe, 21-101nt) or PS modification (PS, 21-101nt) to all nucleotides or all phosphodiester bonds at the invariable part of sgRNA. Complete modification of the invariable part of sgRNA also totally destroyed its function (Fig. 1c). The loss of genome editing ability was further demonstrated by tracking of indels by decomposition (TIDE) analysis and T7EI assays of the amplicon from the GFP locus (Supplementary Fig. 1). These data indicate that the invariable part of the sgRNA cannot be fully modified, and that certain partial modifications can also abolish the function of the sgRNA.

To better identify regions of the sgRNA that might tolerate or be sensitive to chemical modification, we designed patterns of sgRNA chemical modifications guided by the structure of the Cas9–sgRNA complex<sup>26,27</sup>. These structural biology studies demonstrated that the tetraloop and the second stem-loop of sgRNAs extend outside of the Cas9 protein. We hypothesized that these two loops could tolerate modifications in all nucleotides, as long as those modifications did not inhibit loop formation. Indeed, the sgRNA carrying the 2'OMe (2'OMe, Loops) modification at these two loops was functional, and, in fact, it showed a small but significant increase of editing efficiency (89.3% ± 0.3% of GFP<sup>-</sup> cells,  $P < 0.001$ ; Fig. 1d). Notably, incorporating 2'F-modified RNA into these two loops significantly decreased the activity of sgRNA in cells (42.4% ± 0.5% of GFP<sup>-</sup> cells;  $P < 0.001$ ) (Fig. 1d). It is possible that the 2'F modification within loops influences their secondary structure<sup>28</sup>. For this reason, we focused on the 2'OMe modification for the invariable part of sgRNA during the following studies.

Further review of the structure of the Cas9–sgRNA complex<sup>26,27</sup> showed that ~20 nucleotides in the invariable region of sgRNA interact with the Cas9 protein at the 2' hydroxyl (OH) group, mostly through hydrogen bonds (Fig. 1a and Supplementary Table 1). We hypothesized that if we avoided modifying these ‘interacting’ 2'OH groups, and instead modified the remaining nucleotides that did not interact with Cas9 protein at the 2'OH with 2'OMe, then this structure-guided (SG) chemical modification pattern would retain the genome editing activity of sgRNA. Strikingly, this comprehensive, heavily modified sgRNA (named as SG-2'OMe), in which 60 out of 81 nucleotides were modified with 2'OMe RNA, fully retained its activity in cells with a moderate increase in editing efficiency (Fig. 1e). In contrast, a 2'OMe modification at all nucleotides of the invariable region or selectively at the ‘interacting’ 2'OH groups (Reverse-SG-2'OMe) fully abolished the genome editing activity of sgRNA (Fig. 1c,e).

Next, we sought to combine two different chemical modifications, the 2'OMe and the PS. Because the PS modification of all phosphodiester bonds at the invariable part abolished the activity of sgRNA (Fig. 1c), we decided to modify part of the tail region, which has minimal

interaction with Cas9 protein (PS\_70-101nt). This partial PS modification maintained the sgRNA's activity, with a slight increase in activity over the native strand (Fig. 1e). We then combined the patterns of PS (PS\_70-101nt) with 2'OH modification (SG-2'OMe). The combined modification pattern (SG-2'OMe-PS) on the sgRNA significantly enhanced genome editing in cells ( $96.2\% \pm 0.5\%$  GFP<sup>-</sup> cells;  $P < 0.001$ ).

### Structure-guided chemical modification of guide sequences

To explore the rules of chemical modifications at the variable part (the guide sequences), we used a crRNA/tracrRNA system (Fig. 2a) to reduce synthetic cost. We first measured the efficiency of crRNA modified at all nucleotides in the guide sequence with the 2'OMe modification (2'OMe\_1-20nt). We found that full-length modification at the guide sequence with 2'OMe abolished its function (Fig. 2b). We next sought to explore modification of half of the nucleotides in the guide sequence. We modified 1-10 nt or 11-20 nt, as the tail or seed region, respectively (Fig. 2a). We found that the modification of all nucleotides in the seed region abolished the activity of crRNA in cells (Fig. 2b) and the modification of all nucleotides in the tail region also decreased its editing efficiency (Fig. 2b). Based on the structure of sgRNA-Cas9 in the region of guide sequence (Fig. 2a), we identified five nucleotides that interact with Cas9 at their 2'OH group. Among them, four sites are located at the seed region and one at the tail region (Supplementary Table 1). Therefore, we designed two crRNAs, both with a modification of five nucleotides at the seed region and modification of another five nucleotides at the tail region. The first design avoided modifying these 2'OH interacting sites (SG-2'OMe\_1-20nt), whereas the second design intentionally modified all the interacting sites (Reverse-SG-2'OMe\_1-20nt) (Fig. 2a). We found that crRNA SG-1-20nt-2'OMe, which contained 10 2'OMe-modified nucleotides but avoided the modification of the interacting sites, fully retained its activity (Fig. 2b). By contrast, modification of the interacting 2'OH abolished the activity of crRNA in cells (Fig. 2b).

To examine whether 2'F RNA modification on the guide sequence displayed a similar function, we synthesized crRNA with the same designed patterns, but replaced 2'OMe with 2'F. The 2'F modification on the guide sequence followed a similar pattern as the 2'OMe modification (Fig. 2c). The full-length modification of 2'F (2'F\_1-20nt) or "half modification" on the seed region (2'F\_11-20nt) significantly decreased, but did not abolish, its function as happened with the 2'OMe modification. We noted that the physical size of F is smaller than OMe, and there are substantial electronic differences between F and OMe as well, for example, F, but not OMe, can serve as a very weak hydrogen bond acceptor<sup>19,29</sup>, thus, it is possible that modification by 2'F is more compatible with Cas9 binding than 2'OMe.

To identify whether modifications of 2'OH groups and phosphate in the guide sequence can be combined, we first modified the guide sequence with PS only. We found that modification of all 20 nucleotides with PS significantly decreased the activity of crRNA (PS\_1-20nt) (Fig. 2d). Based on the crystal structure, a number of sgRNA phosphates interact with the Cas9 protein (Supplementary Table 2). We therefore sought to avoid modifying these phosphates while replacing others with PS in the guide sequence. This crRNA (SG-PS\_1-20nt)

exhibited significantly ( $P < 0.05$ ) higher potency of genome editing than unmodified crRNA (Fig. 2d).

We then combined the PS with the 2'OMe or the 2'F pattern of modifications described above (Fig. 2e). The combination of the 2'F and PS (SG-PS-2'F) significantly ( $P < 0.05$ ) increased the activity of crRNA, but the combination of the 2'OMe and PS at the guide region (SG-PS-2'OMe) was not tolerated (Fig. 2e). Thus, we decided to combine the 2'F and PS on the guide sequences for further studies. Because this design did not fully modify the 5' end of the crRNA, to prevent cleavage by exonucleases, the first three nucleotides at the 5' end were modified with 2' OMe and PS, or 2'F and PS, in combination with the pattern of SG-PS-2'F\_1-20nt in the guide region. We found the combination (5'-PS-2'OMe\_SG-PS-2'F) resulted in slightly higher editing efficiency than the second combination (5'-PS-2'F\_SG-PS-2'F) (Fig. 2e), thus, we decided to use the first pattern for the guide sequence.

To explore whether such a pattern of modification (5'-PS-2'OMe\_SG-PS-2'F; Fig. 2e) can be used for guide sequences targeting other genes, we synthesized additional crRNA sequences, targeting human genes *HBB* (hemoglobin subunit beta) and *EXM1* (empty spiracles homeobox 1), with and without the above developed chemical modification pattern on the guide sequences (Fig. 2f). We found that this pattern of modification enhanced the formation of indels 2- to 3-fold compared to unmodified crRNAs (Fig. 2f).

To understand whether modifications of individual nucleotide on the variable part could affect the potency of genome editing, crRNAs with modifications at one or several 2'OH groups in the guide sequences were synthesized and evaluated (Fig. 2g). The substitution of even a single 'interacting' 2'OH with 2'OMe significantly decreased or even abolished the activity of crRNA (Fig. 2g). By contrast, the substitution of three non-'interacting' 2'OH with 2'F significantly increased the editing efficiency of the crRNA (Fig. 2g).

### Evaluating e-sgRNA in cells

Next, we sought to combine modifications at the variable and invariable parts (Fig. 3a). We compared the editing efficiency of (1) unmodified sgRNA-targeting GFP (native sgRNA), (2) a published chemical modification of both the 5' and 3' ends (2'OMe and PS modifications of 3 nt at the 5' and 3' end, respectively, named as 5'&3'-sgRNA)<sup>22</sup>, and (3) the pattern of 'enhanced' chemical modification, generated by combining the patterns of modification shown in Figures 1e and 2e, that we term e-sgRNA. Cas9 mRNA and one of these three sgRNAs were delivered to HEK293 cells expressing GFP (Fig. 3b). This e-sgRNA generated a significantly higher number of indels than the 5'&3'-sgRNA and the native sgRNA (43%, 22% and 20%, respectively). In contrast, as previously discussed, native sgRNA transfected into cells already expressing Cas9 induced potent GFP knockout ( $82.3\% \pm 1.7\%$ ) (Fig. 1c and Supplementary Fig. 1). We hypothesize that the modifications of e-sgRNA facilitate function in part by increasing RNA resistance to degradation, which thereby allows persistence while Cas9 mRNA is translated.

To investigate whether this pattern of sgRNA modification can be applied to target other genes, we synthesized e-sgRNA targeting the human *HBB* gene. When delivered with Cas9

mRNA, unmodified *HBB* sgRNA, 5' & 3' -sgRNA and e-sgRNA generated an undetectable percentage, ~15% and ~50% indels, respectively (Fig. 3c). To investigate whether e-sgRNA altered off-target activity, we measured off-target frequencies of *HBB* sgRNA at three predicted genomic loci<sup>22</sup>. For two of the three predicted sites, no off-target activity was found for any of these sgRNAs. Off-target cleavage at site 1 was observed, as previously reported<sup>22</sup>. However, although the *HBB* e-sgRNA produced higher off-target activity at site 1 (Fig. 3d), the off-target/on-target indel frequency ratios remained similar among all three tested sgRNAs.

Cas9 protein may partially protect bound sgRNA from degradation<sup>22</sup>. To investigate whether higher activity of e-sgRNA is observed when Cas9 protein was presented before delivery of sgRNA, different doses of unmodified RNA, 5' & 3' -sgRNA and e-sgRNA targeting GFP were introduced into HEK293 cells expressing GFP and Cas9. Both e-sgRNA and 5' & 3' -sgRNA exhibited similar activity at relatively high doses (500 ng and 250 ng), and showed higher activity than unmodified sgRNA (Supplementary Fig. 2a). The e-sgRNA generated higher indel frequencies than 5' & 3' -sgRNA at a lower dose. The fact that e-sgRNA and 5' & 3' -sgRNA showed similar activity at higher doses when Cas9 protein was already expressed suggests that the binding of Cas9 protein can protect sgRNA from degradation. We then compared the editing efficiency of Cas9 mRNA and e-sgRNA, and ribonucleoproteins (RNPs) of Cas9 protein and e-sgRNA. These two combinations generated similar frequencies of indels in HEK293 cells (Supplementary Fig. 2b).

### Evaluate e-sgRNA *in vivo*

To explore the *in vivo* application of e-sgRNA, we formulated the GFP-targeting sgRNA into a lipid nanoparticle (LNP) (Fig. 4a)<sup>30</sup>. Editing efficiency in liver tissue was first evaluated in a mouse strain constitutively expressing Cas9 and GFP (Cas9-2A-GFP)<sup>31</sup>. We first compared the efficiency of LNPs formulated with native sgRNA, 5' & 3' -sgRNA and e-sgRNA after intravenous injection. After a single injection, unmodified GFP-sgRNA induced a low indel rate at the GFP locus in the liver tissue (~5%), compared with a modest ~22% for 5' & 3' -sgRNA. By contrast, e-sgRNA treatment resulted in a significantly higher rate of indels (~46%) than native or 5' & 3' -sgRNA (Fig. 4b). To explore whether repeated dosing of LNP-sgRNA could increase editing events *in vivo*, we treated mice with two doses of LNP encapsulated e-sgRNA and found a substantial increase in indel frequency (~77%).

To evaluate the *in vivo* potential of non-viral delivery of e-sgRNA with Cas9 for a therapeutically relevant target, we designed two e-sgRNAs targeting the mouse proprotein convertase subtilisin/kexin type 9 (*Pcsk9*) gene (named as PCSK9-1 and PCSK9-2 e-sgRNAs, respectively), a target for the treatment of familial hypercholesterolemia<sup>32</sup>. We encapsulated Cas9 mRNA and both e-sgRNAs in LNPs (Fig. 4c). Five days after a single intravenous administration of these LNPs, serum Pcsk9 was undetectable (Fig. 4d), and total cholesterol had decreased 35–40% (Fig. 4e). We identified a total of 83% ± 3% editing events in the liver genomic DNA, including small indels, a major genomic deletion induced by two sgRNAs and lower levels of inversion (Fig. 4f). A 55-bp gel shift was clearly shown in the PCR products spanning these two sgRNAs (Supplementary Fig. 3).



To investigate whether genome editing occurs primarily in hepatocytes or in other cellular populations in the mouse liver, we perfused and digested the livers, isolated hepatic nonparenchymal cells (NPC) and enriched the hepatocyte population<sup>33</sup>. We found indels in total liver tissue and in the enriched hepatocyte population, but not in NPCs (Supplementary Fig. 4). These data are consistent with previous observations that the LNP delivery system used in this study is specialized for hepatocytes<sup>30</sup>.

We next compared the efficiency of unmodified, 5'&3' and e-sgRNA *in vivo* PCSK9 editing. Unmodified sgRNA of PCSK9-1 or PCSK9-2 guide sequence resulted in ~5% and ~4% indel formations, and 5'&3'-sgRNA generated ~25% and ~18% indels, respectively (Fig. 4g,h). In contrast, injection of one e-sgRNA targeting *Pcsk9* with Cas9 mRNA resulted in ~65% and ~45% indel formations in the mouse liver, respectively (Fig. 4g,h). Furthermore, injection of two e-sgRNAs targeting *Pcsk9* also induced substantially higher indel frequencies in the mouse liver than treatment with two unmodified or 5'&3'-sgRNAs (Supplementary Fig. 5).

To further evaluate the *in vivo* potency of e-sgRNA, we synthesized native, 5'&3' and e-sgRNAs targeting mouse fumarylacetoacetate hydrolase (*Fah*) and *ROSA26* loci. Previously, we have demonstrated efficient LNP-mediated *in vivo* Cas9 mRNA delivery<sup>12</sup>, but we did not observe substantial indel formation in the mouse liver by co-delivery of Cas9 mRNA and an unmodified *Fah* sgRNA in FAH<sup>mut/mut</sup> mice (Supplementary Fig. 6a,b). In contrast, e-sgRNA generated >40% indel formations at the *Fah* locus in FAH<sup>mut/mut</sup> mice, significantly higher than either unmodified or 5'&3'-sgRNA (~14%) (Fig. 4i). Notably, for the same *Fah* guide sequence, the number of indels generated by e-sgRNA is substantially higher than sgRNA expressed by AAV (~24%)<sup>12</sup>. The potency of sgRNAs targeting *ROSA26* was evaluated in wild-type mice. e-sgRNA targeting *ROSA26* also generated significantly ( $P < 0.05$ ) higher levels of indels in the mouse liver than unmodified and 5'&3'-sgRNA (Fig. 4j).

Mice treated with LNP-Cas9 mRNA and *Pcsk9* e-sgRNA showed normal liver histology, body weight and serum biochemistry 5 d after dosing (Supplementary Fig. 7a-c). In addition, histopathology analysis (Supplementary Fig. 7d) and serum biochemistry (Supplementary Fig. 7e,f) performed 24 h and 18 d after dosing appeared normal, suggesting no induction of acute or chronic liver toxicity after gene editing. We measured cytokine arrays 24 h after dosing. Whereas most cytokines were not induced after LNP treatment, interleukin (IL)-6 and G-CSF were induced three- to sixfold after LNP-Cas9 mRNA and native sgRNA or 5'&3'-sgRNA, but not after e-sgRNA treatment (Supplementary Fig. 7g). This may reflect the property of modified RNA to reduce non-specific immune stimulation<sup>34</sup>. To determine whether tissues other than liver were gene edited or not, we measured gene editing efficiency in lung and spleen after treatment of Cas9 and e-sgRNAs targeting *Pcsk9* or *ROSA26* or *Fah*. We found undetectable levels of gene editing in the lung and spleen (Supplementary Fig. 8a-c), indicating the liver specificity of the LNPs used in this study.

We performed GUIDE-Seq<sup>35</sup> to systematically compare the off-target activity of e-sgRNA programmed SpCas9 relative to other guide systems. Genome-wide off-target analysis was performed in mouse Hepa1-6 cells stably expressing SpCas9 and transfected with: (1)

plasmid expressing sgRNAs from a U6 promoter, (2) unmodified sgRNA, (3) 5' & 3' -sgRNA or (4) e-sgRNA. Analysis of the off-target peaks recovered for PCSK9-1 guide sequence revealed few off-target sites (1 or 2) in all four treatment groups with similar nuclease activity between the treatment groups (Fig. 5a–c and Supplementary Table 3). For the PCSK9-2 guide sequence, unmodified, 5' & 3' sgRNA and e-sgRNA shared many off-target sites and the total number of these was substantially reduced relative to plasmid sgRNA expression (Fig. 5a–c). In addition, the bulk activity of the e-sgRNA-programmed SpCas9 across these off-target sites was substantially lower than that observed for plasmid sgRNA expression, which is reflected in an improved specificity ratio. Then we performed deep sequencing to measure the mutation frequency of PCSK9-1 and PCSK9-2's off-target sites from GUIDE-seq, using liver samples from LNP-encapsulated Cas9 mRNA and sgRNA-treated mice. We found no detectable off-target effects in Cas9 with native PCSK9-1 sgRNA-, 5' & 3' -sgRNA- or e-sgRNA-treated livers, indicating the high *in vivo* specificity of this guide sequence (Fig. 5d and Supplementary Table 4). We found undetectable indels in all except one off-target site of PCSK9-2. A low frequency of indels (1%) was detected in one off-target site obtained from GUIDE-seq in the livers of e-sgRNA-treated mice, further highlighting the importance of guide sequence selection to minimize off-target effects (Fig. 5d and Supplementary Table 4)<sup>36</sup>. In our previous studies, we performed GUIDE-Seq for *Fah* sgRNA, and found that only one site (OT1) passed the stringent criteria for potential cleavage sites by Cas9 (ref. 12). We analyzed the OT1 site in the liver harvested from mice with *Fah* e-sgRNA and Cas9 treatment, and we found no detectable indels at the OT1 site (Supplementary Fig. 9).

## DISCUSSION

Using the structure of Cas9–sgRNA complex as a guide, we have identified a number of design criteria that were transferable between different sgRNAs, highlighting the significance of the crystal structure to engineer chemically modified sgRNA. By avoiding modification of 2' OH and phosphate groups in the sgRNA that interact with Cas9 protein, chemical modifications can be made extensively, with ~70% of nucleotides chemically modified. This generalized approach may also be applied to direct comprehensive chemical modifications of sgRNA for other Cas9 proteins, for example, saCas9 (ref. 11), or crRNA in CRISPR–Cfp1 (ref. 37). Our data also suggest that those 2' OH groups of sgRNA that interact with Cas9 by hydrogen bonding play a key role in the formation of a functional Cas9–sgRNA complex. This is distinct from the RNAi machinery, which does not require the 2' OH of siRNA to activate the RNase activity of RISC complexes<sup>21</sup>. Further studies will be needed to further optimize the modification pattern and to uncover potential sequence-specific effects.

Non-viral genome editing is particularly attractive in a therapeutic setting, given the potential advantages of non-viral delivery systems, including ease of scale-up, speed of customization, lack of pre-existing immunity, and the possibility for limiting exposure to nuclease, among other items<sup>14,20</sup>. Here we demonstrate that appropriate chemically modified sgRNA enables very efficient *in vivo* (up to >80%), non-viral vector-mediated genome editing. A number of factors complicate accurate quantification: (1) larger-deletion-bearing alleles will possibly amplify more efficiently, although two differently sized PCR

amplicons showed similar deletion of a 55-bp genomic sequence (Supplementary Fig. 3). (2) The LNP used here is largely hepatocyte-specific<sup>30</sup>, but only ~70–80% percent of the cells in a mouse liver are hepatocytes<sup>38</sup>. Since a large portion of hepatocytes are polyploid<sup>39</sup>, we estimate that hepatocytes account for >80% of total genome copies in the liver<sup>38,39</sup> and show that the majority of gene editing occurred in hepatocytes (Supplementary Fig. 4).

The LNP-mediated co-delivery of Cas9 mRNA and e-sgRNA (LNP–CRISPR) successfully depleted a disease-related protein through creating high levels of indels at the corresponding genomic locus, suggesting its potential for disease treatment (Fig. 4). We note that LNP-mediated siRNA therapy has been evaluated in phase 3 trials and the results are positive<sup>20</sup>. Non-viral genome editing with Cas9 has a number of potential advantages compared with siRNA therapies. First among these, is that a single dose of LNP–CRISPR may provide for long-term therapeutic effects or potentially curative therapy, decreasing or likely eliminating the need for repeated injections. Non-viral delivery of Cas9 as mRNA limits the exposure of the genome to Cas9, decreasing the potential for side effects relative to viral systems where the Cas9 gene is present for months or longer. A number of diseases demand correction or knock-in of sequences<sup>12</sup>. The applications of e-sgRNA to HDR and/or homology-independent targeted integration (HITI) are worth further investigations<sup>40</sup>. Beyond LNP delivery, we believe that these same types of modifications may facilitate delivery of e-sgRNA as a ligand-conjugate form without the need for encapsulation<sup>20</sup>.

Besides the potential therapeutic applications of e-sgRNA, we anticipate that the highly modified sgRNA could be integrated into a range of CRISPR-associated technologies such as CRISPR-mediated imaging, activation and inhibition of targeted genes<sup>41</sup>. We believe that the ability to use nanoparticles to permanently modify the genome of living animals opens the door to a range of therapeutic and industrial applications, and further advances the utility of the Cas9 genome editing system.

## ONLINE METHODS

### RNA synthesis

The sgRNA and crRNA oligos were synthesized by Axolabs (Kulmbach, Germany) using a solid phase synthesis and phosphoramidite chemistry. All RNA oligos were purified by reversed phase high-performance liquid chromatography (HPLC) and subsequently bio-analyzed by liquid chromatography–mass spectrometry (LC-MS). The purity of the final RNA product is approximately 90%. The sequences of all RNAs used are shown in Supplementary Table 5. The *Pcsk9* sgRNA sequences were designed according to sgRNA designer software published by the Broad Institute. Other guide sequences were published elsewhere<sup>12,22,24</sup>.

### Animal experiments

All animal experiments were carried out under the guideline of the MIT and UMass Medical School Animal Care and Use Committee. 1.2 mg/kg nano.Cas9 mRNA and 0.5 mg/kg of each nano.sgRNA targeting *Pcsk9* were introduced into 8- to 10-week-old female C57/BL6 mice (Charles River Laboratories) via tail vein injection. One or two doses of 0.5 mg/kg

nano.sgRNA targeting GFP were intravenously injected into Cas9-P2A-GFP mice<sup>31</sup>. Eleven- to 14-week-old male or female FAH<sup>mut/mut</sup> mice were treated with 0.6 mg/kg nano.Cas9 mRNA and 0.4 mg/kg nano.sgRNA targeting *Fah*.

### Lipid nanoparticles formulation

Cas9 mRNA (Trilink Biotech) and the sgRNA were formulated with ionizable lipid cKK-E12 (ref. 30), cholesterol, C14-PEG 2000, DOPE (1,2-dioleoyl-sn-glycero-3-phosphoethanolamine) in a molecular ratio of 35:46.5:2.5:16, and a cKK-E12:RNA weight ratio of 10:1 using a microfluidic device as previously described<sup>42,43</sup>.

### Cell culture and off-target analysis

HEK293T cells were infected by lentivirus to stably express EF1a-GFP-PGK-Puro and EFs-spCas9-Blast. Cells were transfected with an sgRNA targeting GFP using lipofectamine (Thermo Fisher Scientific). GFP<sup>-</sup> cells were counted by FACS 6–7 d after sgRNA transfection. Off-target sites were predicted using <http://crispr.mit.edu/>. Deep sequencing libraries were made from purified PCR product of predicted genomic locus via Nextera XT kits (Illumina). Libraries were sequenced on Illumina Miseq (150 bp paired-end) or Nextseq (75 bp, paired-end). Reads were mapped to reference sequences and analyzed as previously described<sup>12,44</sup>.

### Determining allele modification frequencies via TIDE analysis, T7EI assay and deep sequencing

Genomic DNA was extracted from cells using QuickExtract- DNA Extraction Solution (Epicentre). Mouse livers, spleens and lungs were ground, and aliquots of tissue powder were taken for extracting genomic DNA using High Pure PCR Product Purification Kit (Roche). PCR was performed to generate the amplicons spanning the targeted sites of sgRNAs. The sequences of primer pairs are shown in Supplementary Table 6. For the Tracking of Indels by Decomposition (TIDE)<sup>45</sup>, the purified PCR products (Qiagen) were sequenced by Sanger method and subsequently analyzed by the TIDE software <http://tide.nki.nl>. For T7EI assay, 200 ng of purified PCR amplicons were denatured, re-annealed and digested with T7 Endonuclease I (New England BioLabs). Digested samples were resolved by electrophoresis in a TBE gel, stained with ethidium bromide, and visualized by UV light. Full-length gels and blots are presented in Supplementary Figure 10. To determine the frequency of events of gene editing after the treatment of Cas9 and two sgRNAs targeting *Pcsk9*, a pair of PCR primers flanking the Cas9 targeting sites of both guide sequences was used to amplify the edited locus. The deep sequencing analysis for determining on-target frequencies is described as above.

### Liver histology, enzyme-linked immunosorbent assays (ELISAs) and serum analysis

Mice were euthanized by CO<sub>2</sub>. The liver of each mouse was taken and fixed in 4% paraformaldehyde overnight. Fixed tissue was embedded in paraffin, sectioned and stained with hematoxylin and eosin (H&E). Blood was collected by retro-orbital puncture to separate serum 5 or 10 d after injection of LNP. ELISA was performed on mouse serum

samples using Pcsk9 ELISA Kit (LifeSpan BioSciences). Serum cytokines, ALT, AST and bilirubin levels were measured as previously described<sup>12,46</sup>.

### Isolation of hepatocytes enriched population and hepatic nonparenchymal cells

The mouse livers were perfused with 1× Hanks buffer briefly followed by digestion buffer containing collagenase<sup>47</sup>. Liver tissues were disrupted gently to form single-cell suspensions, which are filtered through a 100-µm cell strainer. Filtered cells were centrifuged at 30g to pellet hepatocytes. The supernatants were fractionated using Nycodenz (Sigma-Aldrich) to yield hepatic nonparenchymal cells (NPC)<sup>47</sup>.

### GUIDE-seq off-target analysis

We performed GUIDE-seq<sup>35</sup> with some modifications as described previously<sup>44</sup>. Briefly, Hepa1-6 cells stably expressing SpCas9 were transfected using Lipofectamine 3000 transfection reagent (Invitrogen) according to the manufacturer's suggested protocol. For sgRNAs expressed from a U6 promoter, 500 ng of PCSK9-1 or PCSK9-2 plasmid expression vector, 250 ng of a mCherry expression plasmid and 25 pmol of annealed GUIDE-seq oligonucleotide were transfected into cells into each well of a 6-well plate. For transfections of unmodified sgRNA, 5' & 3' sgRNA or e-sgRNAs, 500 ng of sgRNA and 100 pmol of annealed GUIDE-seq oligonucleotide were transfected into cells into each well of a 24-well plate. 48 h after transfection, genomic DNA was extracted with a DNeasy Blood and Tissue kit (Qiagen) according to the manufacturer's suggested protocol. Library preparations were done with original adaptors according to protocols described by Joung and colleagues<sup>35</sup> except that the amount of input genomic DNA was scaled fourfold for the plasmid sgRNA expression samples. Each library was barcoded within the P5 and P7 adaptors for pooled sequencing. The barcoded, purified libraries were deep-sequenced as a pool using a paired-end 150-bp Illumina MiSeq run.

Deep-sequencing data from the GUIDE-seq experiment was analyzed using the GUIDESeq v1.7.2 Bioconductor Package<sup>48</sup> after barcode deconvolution, adaptor removal and alignment to mouse genome (mm10). The default settings for SpCas9 are used except that min.reads is set to 2 and min.peak. score.1strandOnly is set to 2. The potential off-target sites identified for each nuclease are listed in Supplementary Table 3. The Specificity Ratio is calculated as the sum of the unique GUIDE-seq reads at the target site divided by all of the unique reads at all of the off-target sites.

### Statistics

Student's *t*-tests (two-tailed) or One-Way ANOVA with a post-test (Bonferroni post hoc) were applied to determine *P* values by Prism 6 (GraphPad). Significant *P* values are <0.05. The specific statistical method applied, description of replicates can be found in the figure legends. Replicates are defined as numbers of cell cultures for *in vitro* experiments and number of mice for *in vivo* experiments.

## Data availability

The authors declare that the data supporting the findings of this study are available within the paper and its supplementary information files. The original deep sequencing data are available at BioProject: PRJNA355186.

A **Life Sciences Reporting Summary** for this study is available online.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

We thank T. Jacks, P. Sharp, T. Tammela, Z. Weng, G. Gao, E. Sontheimer and A. Vegas for discussions and for sharing reagents, Y. Li and A. Park for technical assistance, and K. Cormier for histology. This work is supported by grants from the National Institutes of Health (NIH), 5R00CA169512, DP2HL137167 and P01HL131471 (to W.X.). V.K. acknowledges support from the Russian Scientific Fund, Grant number 14-34-00017. H.Y. is supported by Skoltech Center and 5-U54-CA151884-04 (NIH). This work is supported in part by Cancer Center Support (core) grant P30-CA14051 from the NIH. We thank the Swanson Biotechnology Center for technical support.

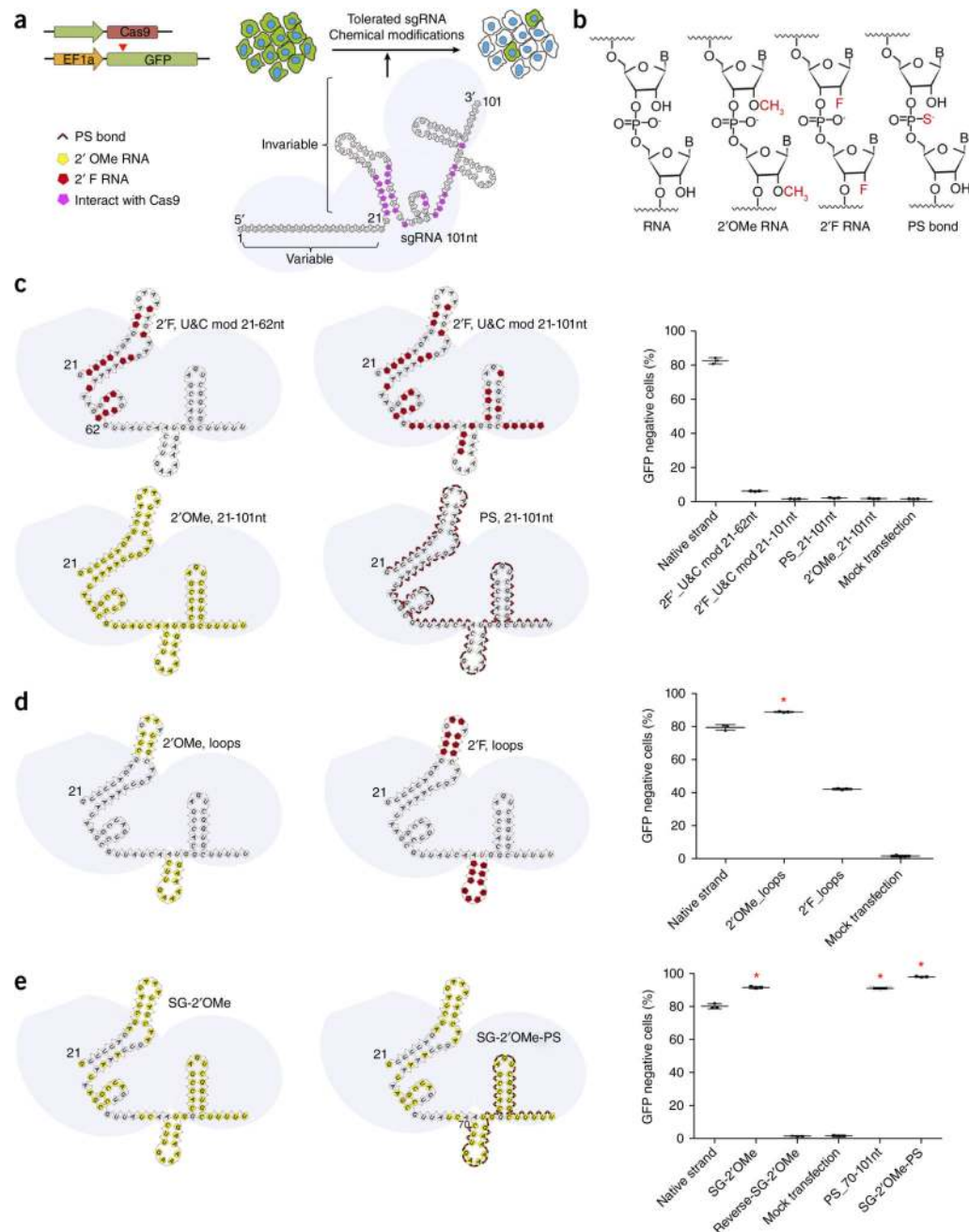
## References

1. Cox DB, Platt RJ, Zhang F. Therapeutic genome editing: prospects and challenges. *Nat. Med.* 2015; 21:121–131. [PubMed: 25654603]
2. Jinek M, et al. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science.* 2012; 337:816–821. [PubMed: 22745249]
3. Cong L, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science.* 2013; 339:819–823. [PubMed: 23287718]
4. Mali P, et al. RNA-guided human genome engineering via Cas9. *Science.* 2013; 339:823–826. [PubMed: 23287722]
5. Doudna JA, Charpentier E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science.* 2014; 346:1258096. [PubMed: 25430774]
6. Swiech L, et al. *In vivo* interrogation of gene function in the mammalian brain using CRISPR-Cas9. *Nat. Biotechnol.* 2015; 33:102–106. [PubMed: 25326897]
7. Long C, et al. Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science.* 2016; 351:400–403. [PubMed: 26721683]
8. Nelson CE, et al. *In vivo* genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science.* 2016; 351:403–407. [PubMed: 26721684]
9. Tabebordbar M, et al. *In vivo* gene editing in dystrophic mouse muscle and muscle stem cells. *Science.* 2016; 351:407–411. [PubMed: 26721686]
10. Schumann K, et al. Generation of knock-in primary human T cells using Cas9 ribonucleoproteins. *Proc. Natl. Acad. Sci. USA.* 2015; 112:10437–10442. [PubMed: 26216948]
11. Ran FA, et al. *In vivo* genome editing using *Staphylococcus aureus* Cas9. *Nature.* 2015; 520:186–191. [PubMed: 25830891]
12. Yin H, et al. Therapeutic genome editing by combined viral and non-viral delivery of CRISPR system components *in vivo*. *Nat. Biotechnol.* 2016; 34:328–333. [PubMed: 26829318]
13. Wang D, et al. Adenovirus-mediated somatic genome editing of Pten by CRISPR/Cas9 in mouse liver in spite of Cas9-specific immune responses. *Hum. Gene Ther.* 2015; 26:432–442. [PubMed: 26086867]
14. Kay MA. State-of-the-art gene-based therapies: the road ahead. *Nat. Rev. Genet.* 2011; 12:316–328. [PubMed: 21468099]
15. Chew WL, et al. A multifunctional AAV-CRISPR-Cas9 and its host response. *Nat. Methods.* 2016; 13:868–874. [PubMed: 27595405]

16. Zuris JA, et al. Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing *in vitro* and *in vivo*. *Nat. Biotechnol.* 2015; 33:73–80. [PubMed: 25357182]
17. Wang M, Glass ZA, Xu Q. Non-viral delivery of genome-editing nucleases for gene therapy. *Gene Ther.* 2017; 24:144–150. [PubMed: 27797355]
18. Behlke MA. Chemical modification of siRNAs for *in vivo* use. *Oligonucleotides.* 2008; 18:305–320. [PubMed: 19025401]
19. Deleavey GF, Damha MJ. Designing chemically modified oligonucleotides for targeted gene silencing. *Chem. Biol.* 2012; 19:937–954. [PubMed: 22921062]
20. Yin H, et al. Non-viral vectors for gene-based therapy. *Nat. Rev. Genet.* 2014; 15:541–555. [PubMed: 25022906]
21. Chiu Y-L, Rana TM. rere siRNA function in RNAi: a chemical modification analysis. *RNA.* 2003; 9:1034–1048. [PubMed: 12923253]
22. Hendel A, et al. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nat. Biotechnol.* 2015; 33:985–989. [PubMed: 26121415]
23. Rahdar M, et al. Synthetic CRISPR RNA-Cas9-guided genome editing in human cells. *Proc. Natl. Acad. Sci. USA.* 2015; 112:E7110–E7117. [PubMed: 26589814]
24. Gilbert LA, et al. CRISPR-mediated modular RNA-guided regulation of transcription in eukaryotes. *Cell.* 2013; 154:442–451. [PubMed: 23849981]
25. Burnett JC, Rossi JJ, Tiemann K. Current progress of siRNA/shRNA therapeutics in clinical trials. *Biotechnol. J.* 2011; 6:1130–1146. [PubMed: 21744502]
26. Nishimasu H, et al. Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell.* 2014; 156:935–949. [PubMed: 24529477]
27. Jiang F, Zhou K, Ma L, Gressel S, Doudna JA. STRUCTURAL BIOLOGY. A Cas9-guide RNA complex preorganized for target DNA recognition. *Science.* 2015; 348:1477–1481. [PubMed: 26113724]
28. Patra A, et al. 2'-Fluoro RNA shows increased Watson-Crick H-bonding strength and stacking relative to RNA: evidence from NMR and thermodynamic data. *Angew. Chem. Int. Ed.* 2012; 51:11863–11866.
29. Howard JAK, Hoy VJ, O'Hagan D, Smith GT. How good is fluorine as a hydrogen bond acceptor? *Tetrahedron.* 1996; 52:12613–12622.
30. Dong Y, et al. Lipopeptide nanoparticles for potent and selective siRNA delivery in rodents and nonhuman primates. *Proc. Natl. Acad. Sci. USA.* 2014; 111:3955–3960. [PubMed: 24516150]
31. Platt RJ, et al. CRISPR-Cas9 knockin mice for genome editing and cancer modeling. *Cell.* 2014; 159:440–455. [PubMed: 25263330]
32. Navarese EP, et al. Effects of proprotein convertase subtilisin/kexin type 9 antibodies in adults with hypercholesterolemia: a systematic review and meta-analysis. *Ann. Intern. Med.* 2015; 163:40–51. [PubMed: 25915661]
33. Yin H, Cheng L, Agarwal C, Agarwal R, Ju C. Lactoferrin protects against concanavalin A-induced liver injury in mice. *Liver Int.* 2010; 30:623–632. [PubMed: 20136718]
34. Robbins M, et al. 2'-O-methyl-modified RNAs act as TLR7 antagonists. *Mol. Ther.* 2007; 15:1663–1669. [PubMed: 17579574]
35. Tsai SQ, et al. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* 2015; 33:187–197. [PubMed: 25513782]
36. Fu Y, et al. High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat. Biotechnol.* 2013; 31:822–826. [PubMed: 23792628]
37. Zetsche B, et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell.* 2015; 163:759–771. [PubMed: 26422227]
38. Si-Tayeb K, Lemaigre FP, Duncan SA. Organogenesis and development of the liver. *Dev. Cell.* 2010; 18:175–189. [PubMed: 20159590]
39. Chen HZ, et al. Canonical and atypical E2Fs regulate the mammalian endocycle. *Nat. Cell Biol.* 2012; 14:1192–1202. [PubMed: 23064266]
40. Suzuki K, et al. *In vivo* genome editing via CRISPR/Cas9 mediated homology-independent targeted integration. *Nature.* 2016; 540:144–149. [PubMed: 27851729]

41. Dominguez AA, Lim WA, Qi LS. Beyond editing: repurposing CRISPR-Cas9 for precision genome regulation and interrogation. *Nat. Rev. Mol. Cell Biol.* 2016; 17:5–15. [PubMed: 26670017]
42. Chen D, et al. Rapid discovery of potent siRNA-containing lipid nanoparticles enabled by controlled microfluidic formulation. *J. Am. Chem. Soc.* 2012; 134:6948–6951. [PubMed: 22475086]
43. Kauffman KJ, et al. Optimization of lipid nanoparticle formulations for mRNA delivery in vivo with fractional factorial and definitive screening designs. *Nano Lett.* 2015; 15:7300–7306. [PubMed: 26469188]
44. Bolukbasi MF, et al. DNA-binding-domain fusions enhance the targeting range and precision of Cas9. *Nat. Methods.* 2015; 12:1150–1156. [PubMed: 26480473]
45. Brinkman EK, Chen T, Amendola M, van Steensel B. Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic Acids Res.* 2014; 42:e168. [PubMed: 25300484]
46. Yin H, Cheng L, Langenbach R, Ju C. Prostaglandin I(2) and E(2) mediate the protective effects of cyclooxygenase-2 in a mouse model of immune-mediated liver injury. *Hepatology.* 2007; 45:159–169. [PubMed: 17187424]
47. Yin H, et al. Lactoferrin protects against acetaminophen-induced liver injury in mice. *Hepatology.* 2010; 51:1007–1016. [PubMed: 20099297]
48. Zhu LJ, et al. GUIDEseq: a bioconductor package to analyze GUIDE-Seq datasets for CRISPR-Cas nucleases. *BMC Genomics.* 2017; 18:379. [PubMed: 28506212]





**Figure 1.** Chemical modifications of invariable part of sgRNA. **(a)** HEK293 cells stably expressing both EF1a promoter-GFP and EFs promoter-spCas9 were incubated with a GFP-targeting sgRNA. Cas9-sgRNA-mediated frameshift NHEJ will yield GFP<sup>-</sup> cells. When a pattern of chemical modification is tolerated by Cas9-sgRNA, the % of GFP<sup>-</sup> cells will be retained. The pink highlighted nucleotides in the invariable region of sgRNA that interact with the Cas9 protein at the 2' hydroxyl (OH) group. **(b)** Chemical modifications of RNA used in the study. **(c-e)** Left: Illustration of full or "U" and "C" chemical modification **(c)**, loops modification **(d)** and structure-guided modification **(e)** in the invariable region (Cas9 binding

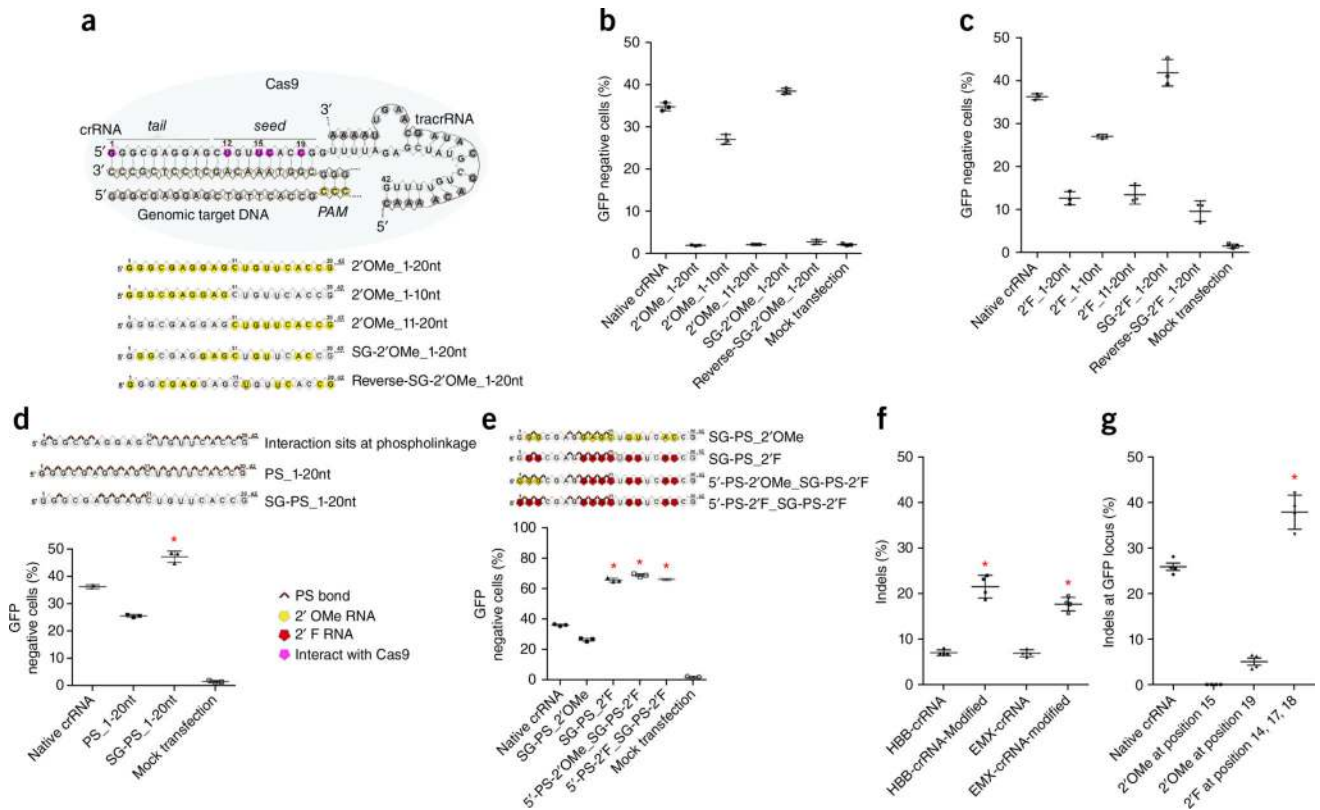
and tail region) of sgRNA. Right: FACS analysis of HEK293 cells described in **a** incubated with sgRNAs with various modifications and without modification (native strand). \* $P < 0.05$ . ( $n = 3$ ), error bars as s.d.

Author Manuscript

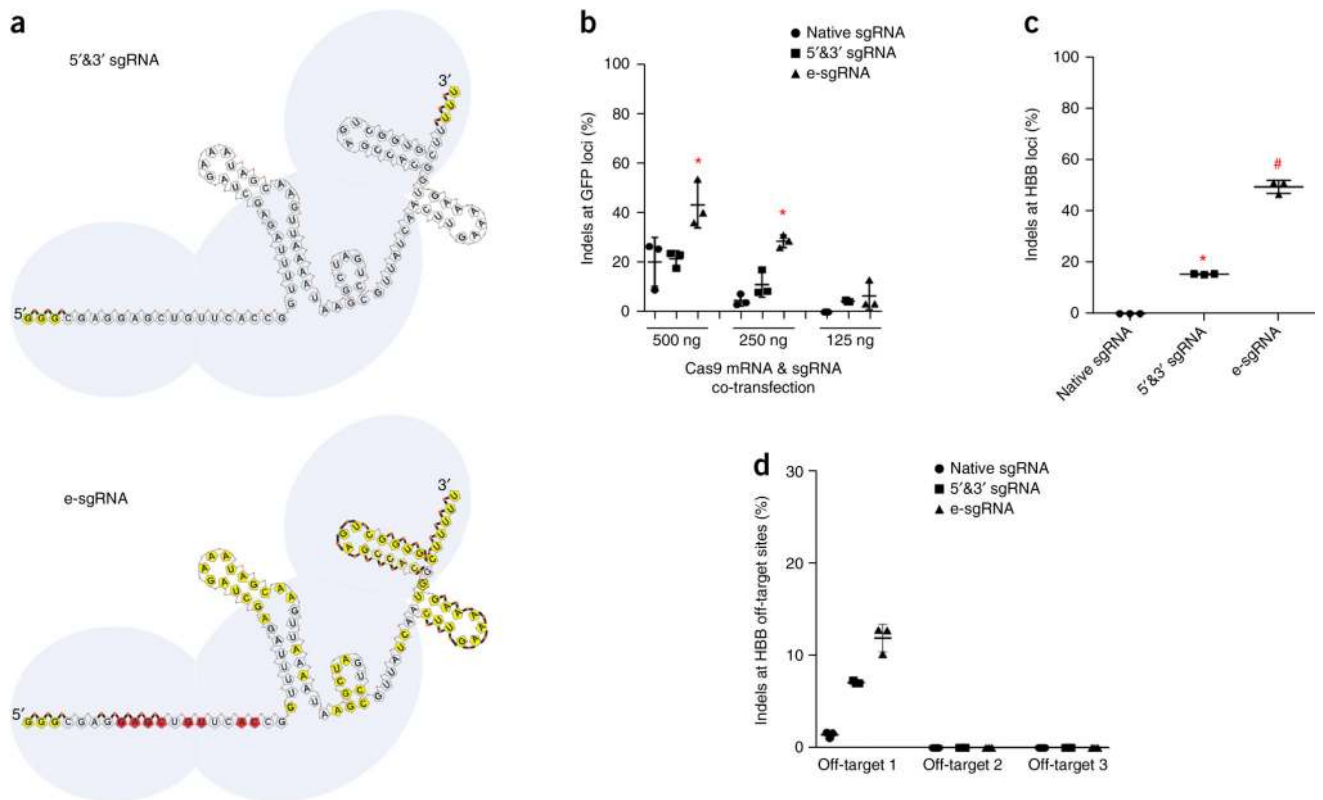
Author Manuscript

Author Manuscript

Author Manuscript

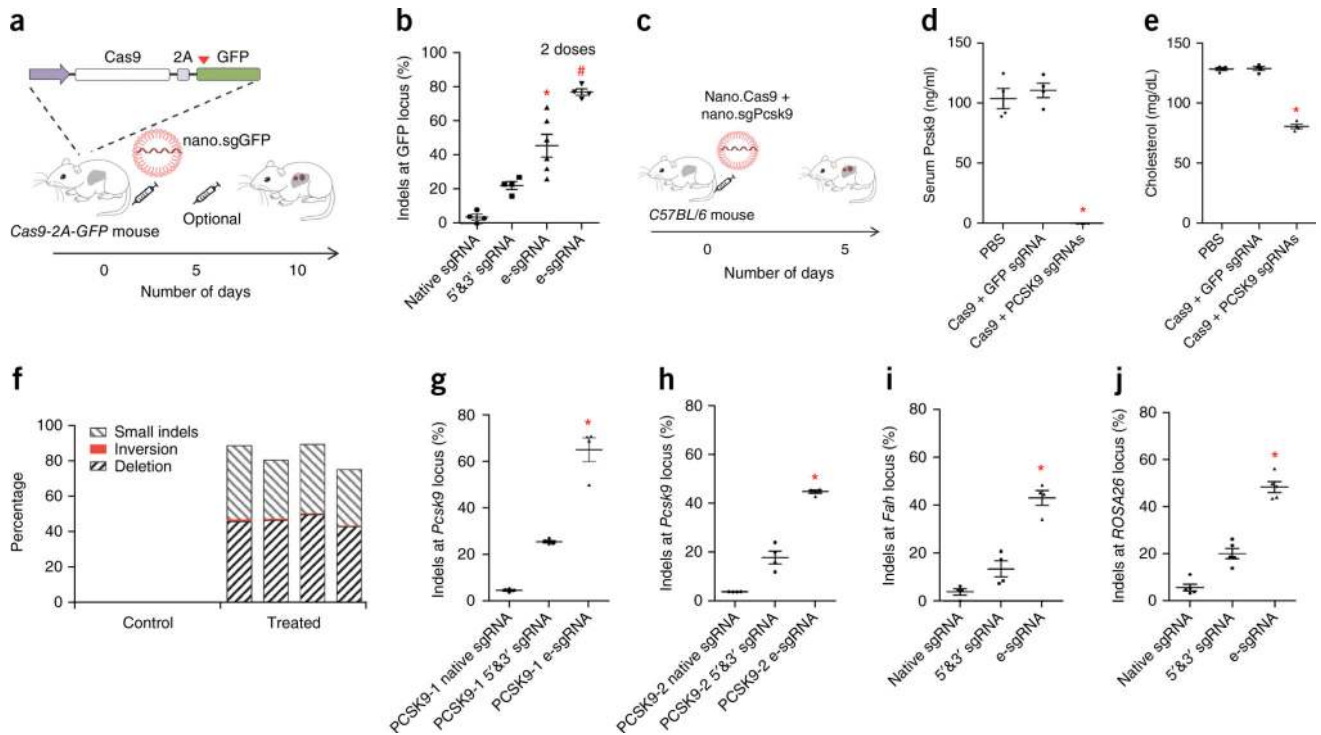
**Figure 2.**

Chemical modifications of guide sequences. (a) crRNAs with 2'OMe modifications and without modifications (native strand). The pink highlighted nucleotides in the guide sequences that interact with Cas9 protein at the 2' OH group. (b–d) crRNA with various patterns of modifications using 2'OMe (b) or 2'F (c) or PS (d). (e) Combination of structure-guided (SG) chemical modification of 2'OMe or 2'F with PS in the guide sequences. HEK293 cells described in Figure 1a were incubated with crRNA in (b–e) and an unmodified tracrRNA. FACS was performed to determine the ratio of GFP<sup>-</sup> cells. (f) crRNAs targeting *HBB* and *EMX-1* were chemically modified with patterns described in e (5'-PS-2'OMe\_SG-PS-2'F), and TIDE analysis was performed to determine indels at the *HBB* and *EMX-1* loci, respectively. (g) crRNAs targeting GFP with modifications of single or multiple 2'OH in the guide region. TIDE analysis was performed to measure indels at GFP locus. \* $P < 0.05$ . ( $n = 4$  in f and g, and  $n = 3$  in others), error bars as s.d.

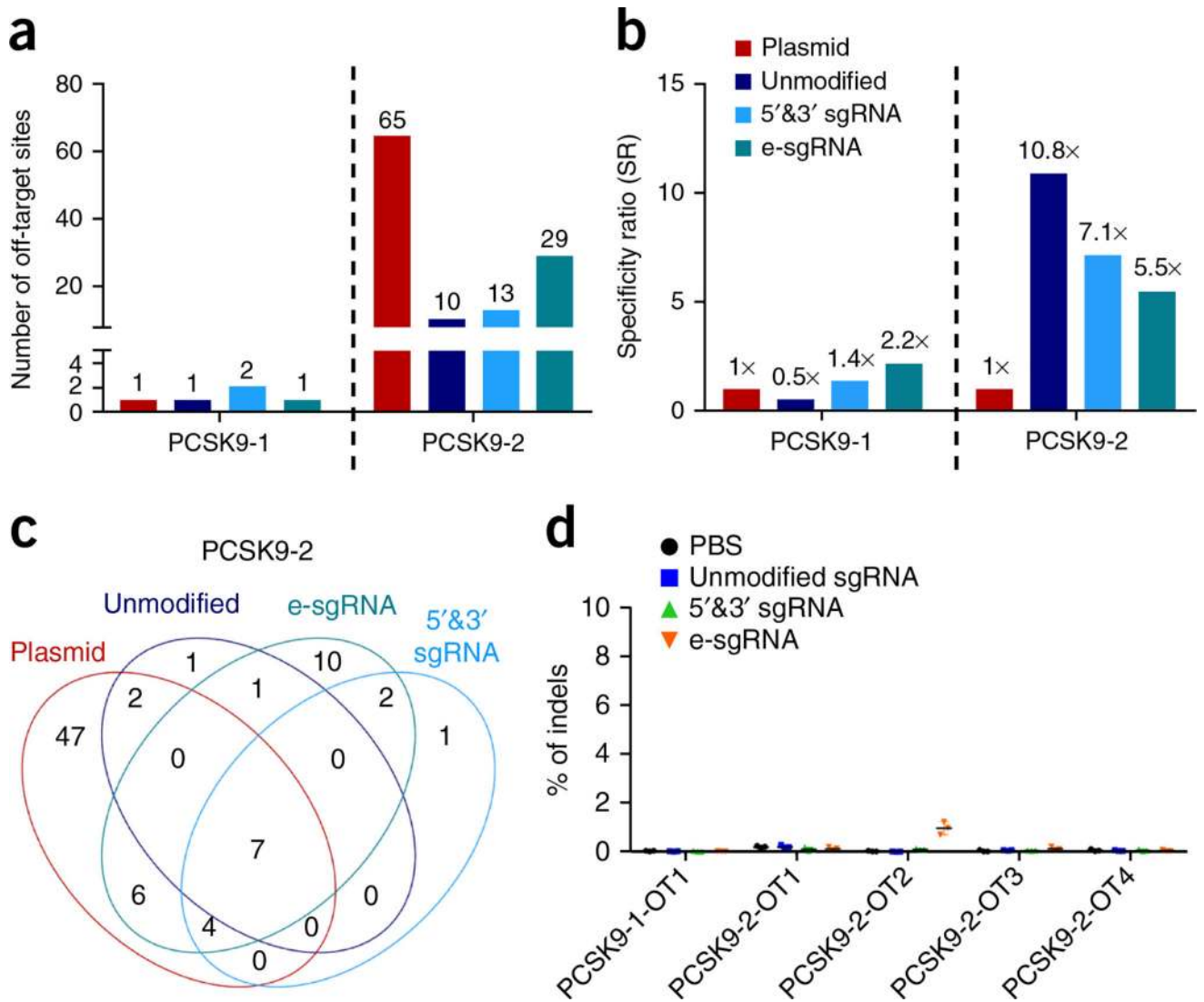


**Figure 3.**

Chemical modifications of sgRNA and its application in human cells. **(a)** Illustration of the conventional 5' and 3' end modification (5'&3'-sgRNA) and our new e-sgRNA design. **(b)** Co-delivery of Cas9 mRNA and sgRNAs targeting GFP into HEK293 cells expressing GFP. FACS analysis was done to determine the number of GFP<sup>-</sup> cells. **(c)** Co-delivery of Cas9 mRNA and sgRNA targeting *HBB* into HEK293 cells. NHEJ events were determined by deep sequencing analysis. **(d)** The editing frequencies of three top off-target sites of *HBB* sgRNA were determined by deep sequencing of amplicon. \* $P < 0.05$ . # $P < 0.05$  compared to the 5'&3' sgRNA treated group ( $n = 3$ ), error bars as s.d.

**Figure 4.**

*In vivo* delivery of chemically modified sgRNAs and Cas9 mRNA induced knockout of targeted gene in the mouse liver. (a) Cas9-2A-GFP transgenic mice were injected with one or two doses of sgRNA encapsulated in lipid nanoparticles (LNP). For mice treated with two doses, the second dose was given 5 d after the first dose. (b) Liver were taken 10 d after first dose. Indels at GFP locus in total DNA from liver were measured by TIDE analysis. (c) C57BL/6 mice were i.v. injected with two e-sgRNAs targeting *Pcsk9* and Cas9 mRNA encapsulated in LNP. (d) The serum Pcsk9 levels. (e) The serum cholesterol levels. (f) The gene editing events at *Pcsk9* locus in total liver DNA, illustrated by deep sequencing. (g–j) C57BL/6 (g,h,j) and *FAH*<sup>mut/mut</sup> (i) mice were i.v. injected with one of native, 5'&3' and e-sgRNA targeting *Pcsk9* (g,h) or *Fah* (i) or *ROSA26* (j) and Cas9 mRNA encapsulated in LNP. Indels at *Pcsk9* (g,h), *Fah* (i) and *ROSA26* (j) loci in total DNA from liver were measured by TIDE analysis. ( $n = 6$  mice in the e-sgRNA/one dose group of b,  $n = 5$  in Fig. 4j and  $n = 4$  mice in others) \* $P < 0.05$ , # $P < 0.05$  compared to the one dose e-sgRNA-treated group, error bars as s.e.m.



**Figure 5.** GUIDE-seq genome-wide off-target analysis of nuclease activity for SpCas9 programmed with PCSK9-1 or PCSK-2 sgRNA expressed from a U6 promoter (plasmid), unmodified, end-modified (5'&3') or e-sgRNA. **(a)** The bar chart indicates the number of off-target peaks detected in the GUIDE-seq data for each type of sgRNA. **(b)** The bar chart indicates the fold improvement in Specificity Ratio (SR) (number of unique reads at the target site/sum of the unique reads at all off-target sites) for each sample relative to the plasmid expressed sgRNA. **(c)** The Venn diagram displays the distribution of the number of GUIDE-seq identified off-target sites that are common or unique for each given treatment group for PCSK-2. For the ten off-target sites unique to the e-sgRNA, nine of these sites have peak scores less than 16, which indicates that they are either weak sites or false positives. For reference the target site score is 9,484 (Supplementary Table 3). **(d)** Off-target sites obtained from GUIDE-seq were amplified using liver samples from mice treated with LNP-encapsulated Cas9 mRNA, PCSK9-1 and PCSK9-2 sgRNA. Deep-seq was performed to

determine mutation frequency. The bars indicate total indel frequencies at the off-target sites.  $n = 3$  mice, error bars as s.e.m.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript