nature communications

Article

https://doi.org/10.1038/s41467-025-58653-1

Modified pegRNAs mitigate scaffold-derived prime editing by-products

Received: 7 August 2024

Accepted: 20 March 2025

Published online: 09 April 2025

Check for updates

Panagiotis Antoniou[1,7], Louis Dacquay[1,2,7], Hanna Mårtensson[1], Katja Madeyski-Bengtson[3], Anna-Lena Loyd[3], Anna Shiriaeva[1,2], Euan Gordon[4], Salman Mustfa[5], George Thom[5], Pei-Pei Hsieh[1], Saša Šviković[1], Mike Firth[6], Nina Akrap[1], Marcello Maresca[1] ✉ & Martin Peterka[1] ✉

Prime editors (PEs) employ reverse transcriptase (RT) to install genomic edits using a template within the prime editing guide RNA (pegRNA). RT creates a 3′ genomic flap containing the intended edit. However, reverse transcription can continue beyond the template, incorporating the pegRNA scaffold sequence into the 3′ flap. These scaffold-derived by-products can be installed alongside the intended edit, reducing prime editing precision. Here, we develop a method that prevents RT from accessing the scaffold, thereby mitigating such by-products. We demonstrate that an internal abasic spacer or 2′-O-methylation within the pegRNAs terminates RT at the end of the template. This prevents scaffold-derived sequences from being incorporated into the target locus. We benchmark these pegRNAs in different cell types and demonstrate that they can be used with processive PEs such as PE6d or PE**. Our findings provide a simple approach to mitigate a common prime editing by-product and improve prime editing precision.

Prime editing is an advanced genome engineering approach that allows installation of various types of small edits with high precision. Prime editor (PE) consists of a Cas9 enzyme fused via a flexible linker to a reverse transcriptase (RT)[1]. PE uses a prime editing guide RNA (pegRNA), which contains an RT template (RTT) encoding the intended edit, and a primer binding sequence (PBS) complementary to the target site. Early steps of prime editing include pairing of the PBS sequence with the cleaved non-target DNA strand and generation of a genomic 3′ flap by reverse transcription of the RTT. Successful annealing and ligation of the edited 3′ flap into the genome results in the incorporation of the desired prime edit into the target site. However, the 3′ flap may contain additional, scaffold-derived sequences caused by the readthrough of RT into the pegRNA scaffold which directly neighbors the RTT. These scaffold-derived sequences have been widely observed co-occurring with intended prime edits at various frequencies and represent a common but unresolved prime editing by-product[1–8]. Moreover, recent advancements in PE variants, which have enhanced processivity, have been shown to increase the levels of scaffold incorporations[6,9].

The structure of PE in the termination state revealed that RT can freely transcribe at least the first two nucleotides of the pegRNA scaffold after which it terminates at the third nucleotide (U94) which is in proximity to Cas9 and likely inaccessible for RT due to steric clash[8]. The additional scaffold-derived sequences in the 3′ flap normally do not anneal to the target site and have increased likelihood to be removed by cellular exonucleases. However, inadvertent partial complementarity with the target site may lead to incorporation of these sequences into the genome. Indeed, recoding of the pegRNA scaffold sequence to avoid complementarity decreased the frequency of scaffold incorporations[2]. After annealing and ligation of the 3′ flap, the

[1]Genome Engineering, Discovery Sciences, R&D, AstraZeneca, Gothenburg, Sweden. [2]Promega Corporation, Madison, WI, USA. [3]Translational Genomics, Discovery Sciences, R&D, AstraZeneca, Gothenburg, Sweden. [4]Protein Science, Structure and Biophysics, Discovery Sciences, R&D, AstraZeneca, Gothenburg, Sweden. [5]RNA Therapy, Discovery Sciences, R&D, AstraZeneca, Cambridge, UK. [6]Data Sciences and Quantitative Biology, Discovery Sciences, R&D, AstraZeneca, Cambridge, UK. [7]These authors contributed equally: Panagiotis Antoniou, Louis Dacquay. ✉e-mail: marcello.maresca@astrazeneca.com; martin.peterka@astrazeneca.com

scaffold-derived mutations can be removed by the mismatch repair pathway (MMR), inhibition of which led to an increase of scaffold incorporations[2].

Previously, we and others have used PEs with the fully active SpCas9 nuclease (referred to here as PEn)[5,10–14]. One such prime editing strategy, termed primed-insertion (PRINS), was tailored to use non-homologous end-joining (NHEJ) to harness the robustness and high activity of this DNA repair pathway across different cell types. During PRINS editing, the 3′ flap, including any additional scaffold-derived nucleotides, is rapidly inserted into the double strand break by NHEJ and thus the scaffold-derived insertions are highly prominent by-products of PRINS editing, hindering its precision[5]. This makes PRINS a suitable system to assess more precisely the extent of scaffold incorporation and to evaluate strategies to improve editing precision.

In this work, we show that RT readthrough into the pegRNA scaffold is prevalent and develop modified pegRNAs that mitigate scaffold-derived prime editing by-products by precisely terminating RT at the end of RTT. Importantly, all of these are common, commercially available RNA oligonucleotide modifications and can be easily applied to varying prime editing systems without any customization of the pegRNA scaffold sequence or PE itself. Finally, we apply these modifications to improve various types and applications of prime editing.

Results

Abasic sites within pegRNAs precisely block reverse transcription by Moloney murine leukemia virus (M-MLV) RT

Abasic sites within oligonucleotide templates block primed extension by various DNA polymerases directly upstream of the abasic site[15]. We thus wondered whether a single abasic site in the pegRNA would block RT and could be used to prevent undesired readthrough into the pegRNA scaffold. To this end, we devised a cellular assay that uses PRINS, a prime editing approach with high levels of pegRNA scaffold incorporation[5]. In this assay, which is based on PEn, a prime editor with the Cas9 nuclease, the 3′ flaps, including those with scaffold-derived sequences caused by RT readthrough, are efficiently captured at the DNA double-strand break (DSB) via NHEJ (Fig. 1a). Modified nucleotides placed directly after RTT can be probed for their ability to block RT and prevent scaffold incorporation events measured by targeted amplicon sequencing (amplicon-seq) (Fig. 1b). To test the effect of abasic sites on prime editing, we designed a synthetic pegRNA to insert a single nucleotide into the human *PCSK9* locus where the one nucleotide long RTT is followed by the pegRNA scaffold (Supplementary Fig. 1a). Utilizing this assay, we edited K562 cells with an unmodified pegRNA, or a pegRNA modified with either riboabasic spacer (rSp-pegRNA), or C3 abasic spacer (C3-pegRNA) positioned between RTT and the pegRNA scaffold (Fig. 1c and Supplementary Fig. 1b). While editing with the unmodified pegRNA led to high levels of

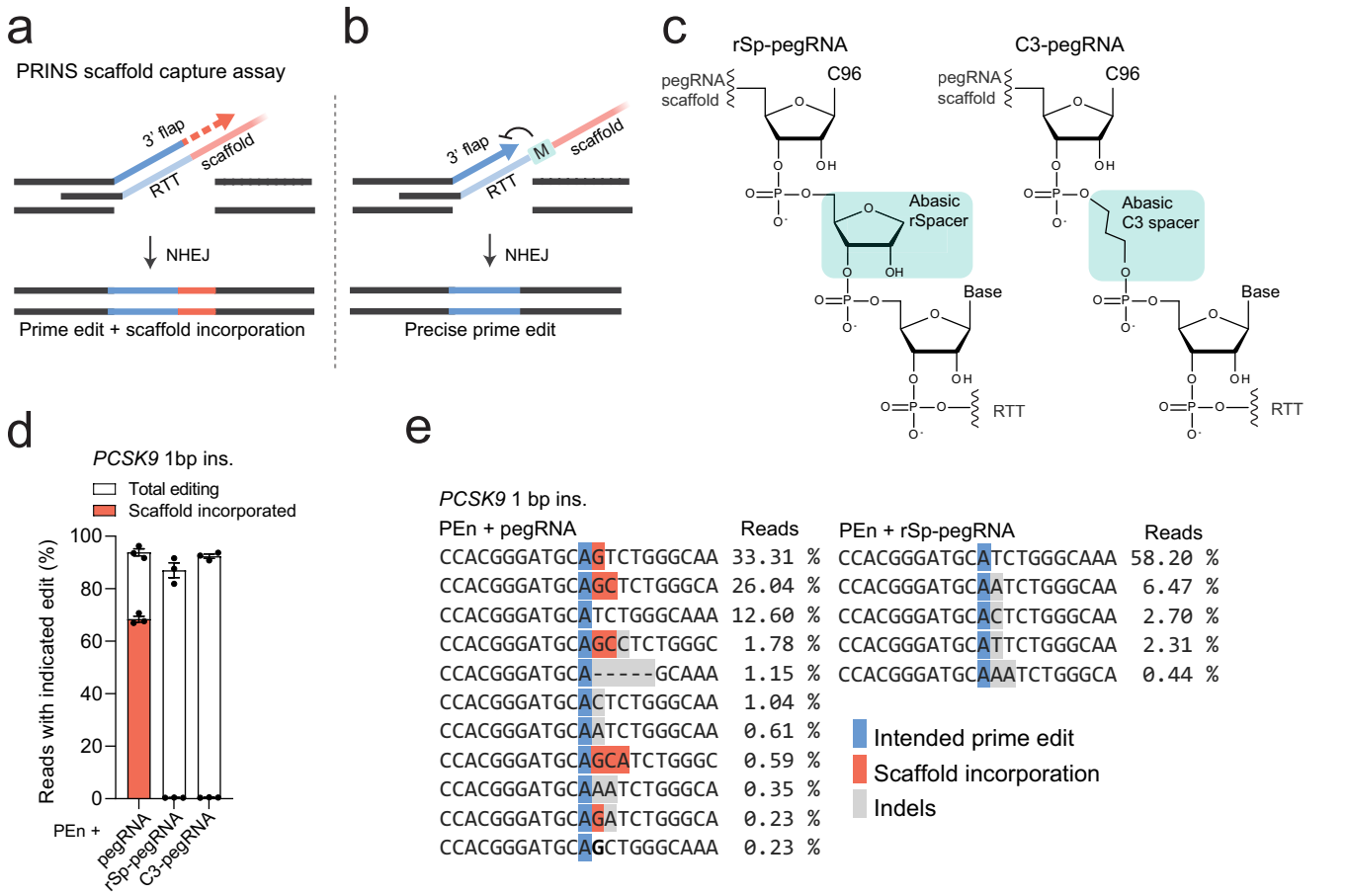


Fig. 1 | Abasic sites within synthetic pegRNAs precisely block reverse transcription by M-MLV RT. a Schematic representation of pegRNA scaffold capture by NHEJ in the PRINS editing assay. PRINS editing results in unintended pegRNA scaffold integrations due to flap extension past the intended reverse transcription template (RTT). **b** PRINS editing using modified (M) pegRNAs to prevent RT from pegRNA scaffold readthrough. **c** Chemical structures of modifications within rSp-pegRNA and C3-pegRNA. C96 = cytosine at the position 96 of the SpCas9 pegRNA scaffold. **d** Scaffold incorporation analysis by amplicon-seq of *PCSK9* editing in K562 cells electroporated with PEn mRNA and unmodified pegRNA, rSp-pegRNA, or C3-pegRNA. Plots show mean ± SD of $n = 3$ biological replicates. **e** Representative CRISPResso2 alignment of PEn + pegRNA and PEn + rSp-pegRNA prime editing outcomes from experiment in Fig. 1d. Source data are provided as a Source Data file.

scaffold incorporations with up to 70% of edits that contained scaffold-derived sequences, both rSp-pegRNA and C3-pegRNA mitigated scaffold incorporations likely due to a robust and precise blocking of M-MLV RT directly upstream of the modified site (Fig. 1d, e). In line with the recently solved PE structure, where the last two nucleotides of SpCas9 pegRNA scaffold (C96 and G95) were shown to be accessible to RT[8], the most abundant scaffold integrations were 1–2 nt long (Fig. 1e and Supplementary Fig. 1a).

To further test the robustness of RT blocking with abasic spacers, we used PRINS to install small insertions (1–8 bp) in four different loci and in four human cell lines (Fig. 2a). Consistent with our initial findings, rSp-pegRNAs and C3-pegRNAs mitigated scaffold readthrough and incorporation, and increased efficiency of precise editing across all targets and cell lines (Fig. 2b). This resulted in significant improvement of PRINS precision that ranged from 1.4- to 6.6-fold increase across all four targets and cell lines (Fig. 2b). Therefore, both rSp-pegRNAs and C3-pegRNAs eliminate scaffold incorporation by blocking M-MLV RT and can be used to increase precision of PRINS editing.

rSp-pegRNAs allow precise and efficient editing of primary human hepatocytes (PHHs) and mouse embryos

Encouraged by these results, we aimed to use the increased precision of PRINS in difficult-to-edit systems. PRINS installs insertions via NHEJ, the predominant DSB repair pathway in most human cell types which is independent on cell cycle progression (Zhao et al., 2020). Hence, we hypothesized that PRINS could be a useful tool for precise editing of postmitotic cells. To test this, we used PRINS to edit PHHs. We transfected PHHs with PEn mRNA and a corresponding synthetic pegRNA or rSp-pegRNA to install small insertions at the *PCSK9* and *HBEGF* loci. Analogous to cell lines, amplicon-seq revealed high overall prime editing with a substantial portion of additional scaffold sequences, which were mitigated using rSp-pegRNA, leading to improved precision as well as efficiency, reaching up to 40% of precise editing at the therapeutically relevant *PCSK9* locus (Fig. 3a).

NHEJ is highly active in mouse embryos[16]. We thus speculated that PRINS might be an effective method for precise genome editing in this system. We electroporated mouse zygotes with PEn and unmodified pegRNA or rSp-pegRNA ribonucleoprotein complexes to install a 5 bp insertion into the endogenous *Map3k15* locus, a therapeutic target for diabetes[17]. The electroporated embryos were cultured in vitro for 5 days after which they were lysed and subjected to amplicon-seq of *Map3k15* which revealed that rSp-pegRNA eliminated scaffold insertions, increased precise prime editing, and increased precision 2-fold compared to editing with unmodified pegRNA (Fig. 3b). In the rSp-pegRNA group, 7 out of 24 embryos analyzed contained precise prime edits in at least 50% of mapped reads, with one reaching up to 100% (Supplementary Fig. 2). These results serve as a basis for further exploration of PRINS as an efficient method to install small insertions in mouse embryos.

Taken together, we show that reverse transcription by PEs can be precisely blocked by an abasic spacer in the pegRNA and this can be used to prevent undesired readthrough beyond RTT into the pegRNA scaffold. Our data also show that pegRNAs with abasic spacers are stable and active in a variety of cell models including primary cells and mouse embryos.

Internal 2′-O-methylation of pegRNAs precisely blocks reverse transcription by M-MLV-RT

To explore additional ways to block RT we also tested the effect of 2′-O-methylation (2′-O-Me) on RT readthrough into the pegRNA scaffold. 2′-O-Me is an RNA backbone modification commonly used to increase stability of synthetic RNA oligonucleotides including gRNAs[18]. 2′-O-Me-modified nucleotides are known to block reverse transcription in vitro, possibly by creating conformational “bumps” on the methylated template[19]. Thus, we speculated that an internal 2′-O-Me modification of the pegRNA could act as an RT block to prevent readthrough into the pegRNA scaffold. We used our PRINS *PCSK9* editing assay (Fig. 1a, b) but instead of introducing abasic spacers into pegRNAs, we designed a pegRNA that contained 2′-O-Me at the cytosine 96 (C96) position of the scaffold (C96Me-pegRNA) (Fig. 4a and Supplementary Fig. 1c). We compared editing with C96Me-pegRNA and unmodified pegRNA in K562 cells. Amplicon-seq showed 5.3-fold decrease of scaffold integration, confirming strong RT blocking immediately upstream of C96 (Fig. 4b and Supplementary Fig. 1c). To test the stringency of RT blocking by 2′-O-Me we edited different cell lines with C96Me-pegRNAs targeting *PCSK9* and *HBEGF* loci. Amplicon-seq analysis demonstrated universally beneficial effects of C96Me-pegRNAs on the prime editing outcomes. For both targets and in all the cell lines tested, the frequency of reads containing scaffold incorporation was significantly reduced and this led to a concomitant increase in the frequency of precise prime edits (Fig. 4c, d). We note that C96Me-pegRNAs did not block scaffold readthrough completely, unlike abasic pegRNAs, which eliminated scaffold integrations below detection limit (Fig. 1d and Supplementary Fig. 3). The readthrough levels appeared to be dependent on target site and cell line (Fig. 4c). The residual scaffold integrations were qualitatively similar, comprising of 1–2 bp long scaffold insertions (Supplementary Fig. 4). We next tested whether multiple methylations could decrease residual scaffold readthrough observed with singly methylated C96Me-pegRNAs. We conducted an experiment to investigate the effect of multiple methylations on RT readthrough. We used PRINS to edit the *Map3k15* locus in murine AML12 cells. We compared PRINS editing with the following pegRNAs: internally unmethylated pegRNA, singly methylated pegRNA (C96Me-pegRNA), doubly methylated pegRNA (G95MeC96Me-pegRNA), and a highly methylated pegRNA (HiMod-pegRNA). The highly methylated design was adapted from Finn et al.[20], where multiple methylations were shown to increase the stability and activity of sgRNAs in vivo. Because 2′-O-Me modifications in this design fully cover 28 nucleotides at the 3′ end of the gRNA scaffold, we reasoned they could serve as a potent stop signal for RT. The results show that the overall rates of scaffold incorporations were largely mitigated by C96Me-pegRNA and were not further reduced with multi-methylated pegRNAs (Fig. 4e). However, we noted a qualitative difference—multiple methylations prevented longer scaffold insertions, suggesting more effective RT termination with multiple consecutive methylations (Supplementary Fig. 5). Therefore, using multiple methylations is advisable in situations where longer scaffold integrations occur at higher frequencies.

In conclusion, although less stringent than abasic sites, our results show that 2′-O-Me robustly blocks reverse transcription by M-MLV RT.

Modified pegRNAs mitigate scaffold incorporation by PE3 and highly processive PE systems

Finally, we sought to apply modified pegRNAs to nickase-based prime editing systems. We used pegRNAs previously shown to induce scaffold incorporations[1,9,21]. First, we used rSp-pegRNA to edit *FANCF* in K562 cells (Fig. 5a). Editing of *FANCF* with the PE3 system and an unmodified pegRNA resulted in ~0.5% of scaffold-containing reads, which represented ~3.4% of prime edited reads (Fig. 5b). rSp-pegRNA decreased the frequency of scaffold-containing reads to 0.01%, which is below the detection limit of amplicon-seq (Fig. 5a), suggesting potent termination of RT at the modified site in the rSp-pegRNA.

Next, to test the 2′-O-Me-mediated RT blocking and its stringency, we used recently engineered PEs that possess higher processivity, PE** and PE6d[6,9,22]. PE** uses an M-MLV RT mutant with increased dNTP affinity to overcome conditions where intracellular dNTP levels are limiting for prime editing[22,23]. Similarly, PE6d contains an evolved M-MLV RT variant with improved processivity beneficial for installation of longer insertions[6]. Both PE systems lead to increased scaffold incorporations, that could be potentially mitigated by C96Me-

https://doi.org/10.1038/s41467-025-58653-1

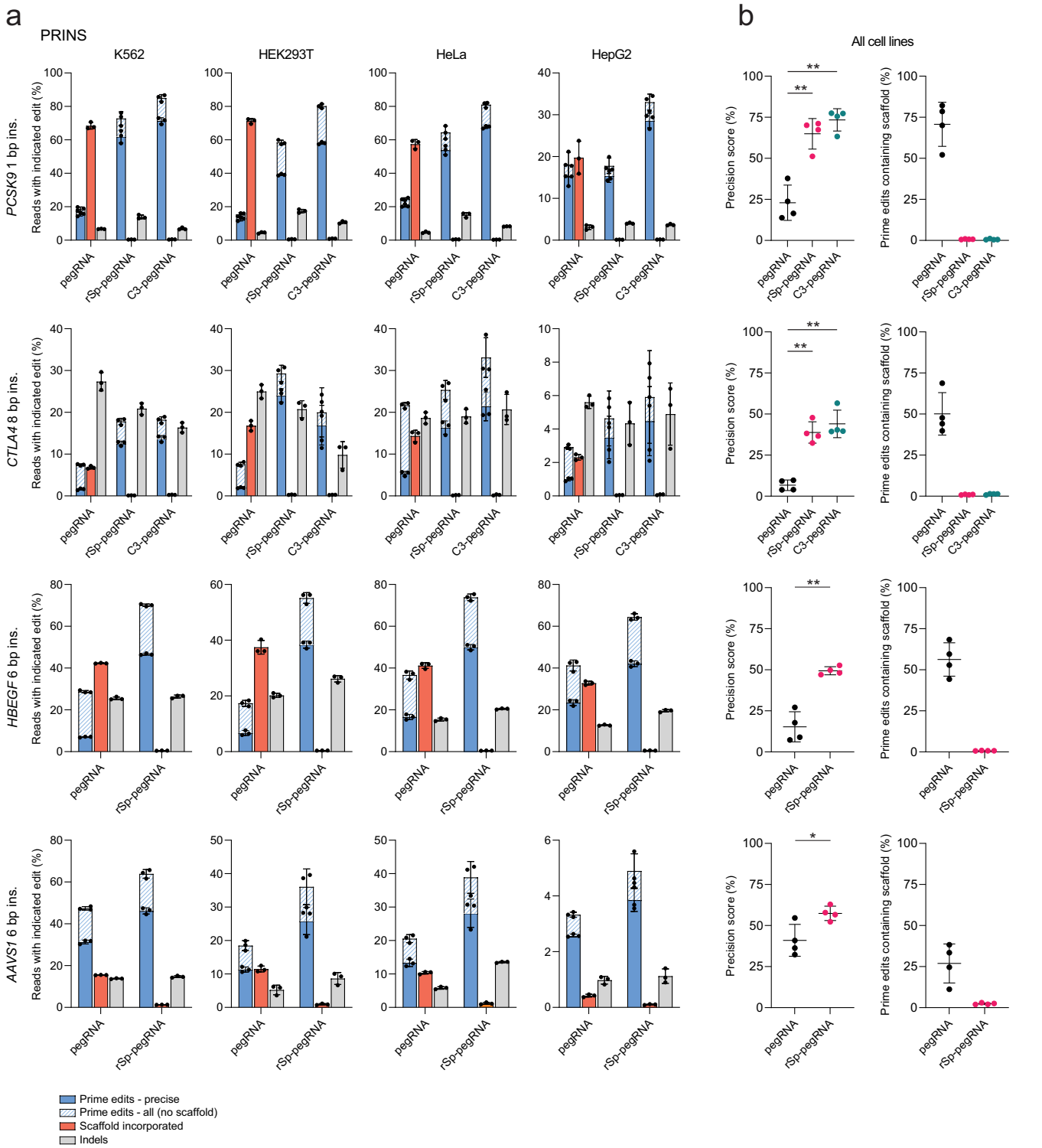
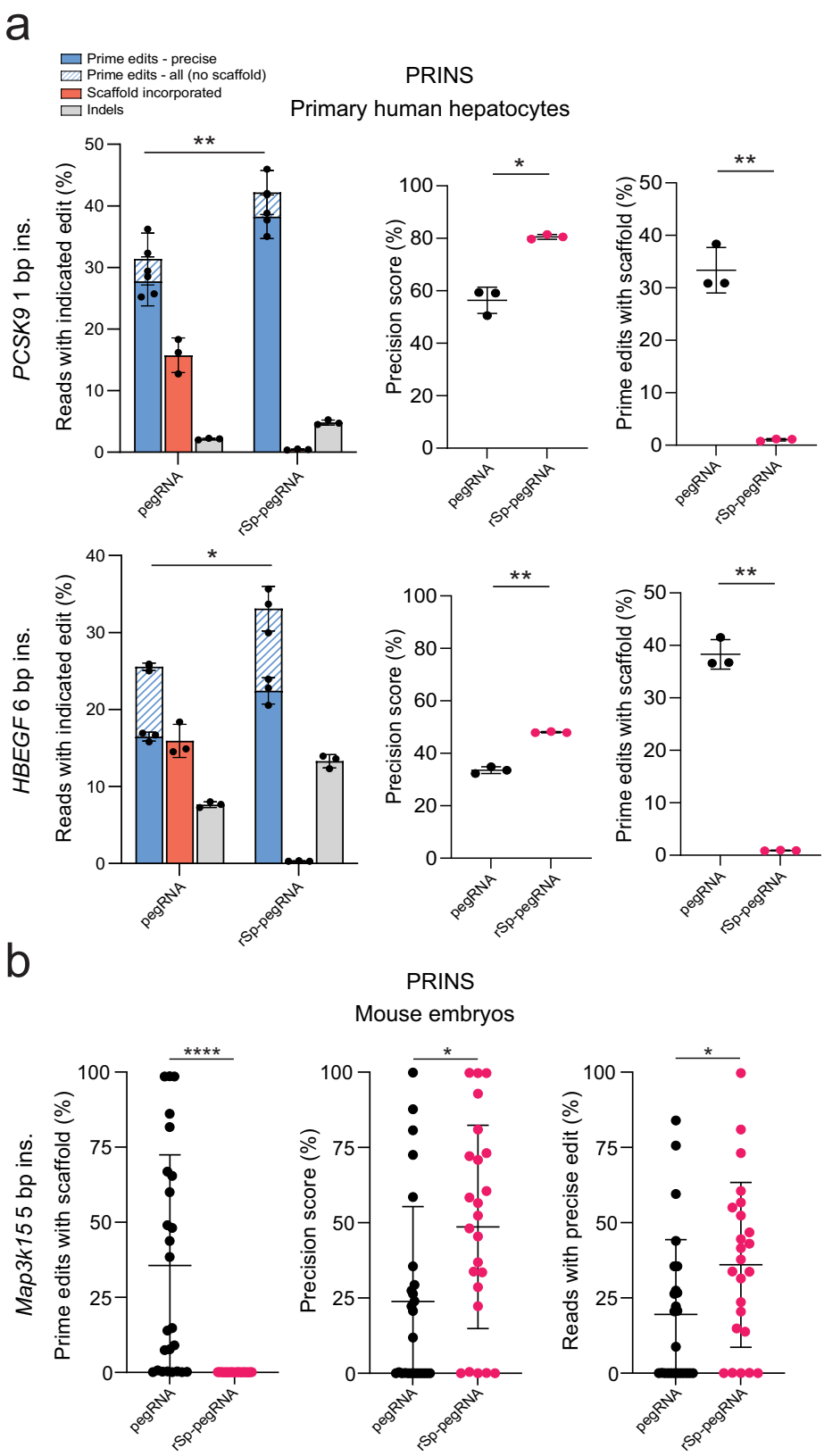


Fig. 2 | PRINS editing with modified pegRNAs in human cell lines. a PRINS editing at *PCSK9*, *CTLA4*, *HBEGF and AAVS1* genomic loci in four different human cell lines (K562, HEK293T, HeLa and HepG2) using electroporation of PEn mRNA in combination with synthetic pegRNA, rSp-pegRNA or C3-pegRNA to install indicated small insertions (ins.). Editing outcomes were analyzed by amplicon-seq and quantified using CRISPResso2 in the prime editing mode. Plots show mean ± SD of $n = 3$ biological replicates. Prime edits precise = intended prime edits; Prime edits all (no scaffold) = precise prime edits + prime edits co-occurring with indels not derived from scaffold; Scaffold incorporated = prime edits with at least one additional nucleotide derived from scaffold; Indels = non-prime edited insertions or deletions. **b** Quantification of data from (**a**). "Precision score" was calculated as total number of amplicon-seq reads with precise prime edit per overall editing. "Prime edits containing scaffold" was calculated as total number of amplicon-seq prime edited reads with scaffold integration per total prime edited reads. Error bars represent mean ± SD. Datapoints represent 4 different cell line experiments in (**a**). Statistical difference was determined using Student's *t* test (paired, two-tailed). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Calculated *P* values: *PCSK9*: pegRNA vs. rSp-pegRNA = 0.0032, pegRNA vs. C3-pegRNA = 0.0014; *CTLA4*: pegRNA vs. rSp-pegRNA = 0.0039, pegRNA vs. C3-pegRNA = 0.0057; *HBEGF*: pegRNA vs. rSp-pegRNA = 0.0034, AAVS1: pegRNA vs. rSp-pegRNA = 0.0154. Source data are provided as a Source Data file.

https://doi.org/10.1038/s41467-025-58653-1



pegRNAs. First, we used Cas9 nuclease variants of PE6d (termed PEn6d) and PE** (termed PEn**) to allow for maximal capture of any potential scaffold-containing 3′ flaps at the DSB. We detected high levels of total reads with scaffold incorporations (20.5% for PEn6d and 13.7% for PEn**; Fig. 5c). Reads with scaffold incorporations corresponded to 37.3% and 26.0% of the total prime edited reads, for PEn6d and PEn**, respectively (Fig. 5d), demonstrating that a large fraction of 3′ flaps contained scaffold-derived sequences. Scaffold incorporations introduced by PEn6d and PEn** were similar and contained long stretches of scaffold-derived sequences, confirming the high processivity of RTs used in these prime editors (Supplementary Fig. 6a). This frequent readthrough past the third nucleotide of the scaffold

Fig. 3 | PRINS editing with modified pegRNAs in primary human hepatocytes and mouse embryos. a PRINS editing at *PCSK9* and *HBEGF* genomic loci in primary human hepatocytes transfected with PEn mRNA and a synthetic pegRNA or rSp-pegRNA to install small insertions (ins.). Editing outcomes were analyzed by amplicon-seq and CRISPResso2 in the prime editing mode. "Precision score" was calculated as total number of amplicon-seq reads with precise prime edit per overall editing. "Prime edits with scaffold" was calculated as total number of amplicon-seq prime edited reads with scaffold integration per total prime edited reads. Plots show mean ± SD of $n = 3$ biological replicates. Statistical difference was determined using Student's t test (paired, two-tailed). Prime edits precise = intended prime edits; Prime edits all (no scaffold) = precise prime edits + prime edits co-occurring with indels not derived from scaffold; Scaffold incorporated = prime edits with at least one additional nucleotide derived from scaffold; Indels = non-prime edited insertions or deletions; $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Calculated P values: *PCSK9* "Prime edits - precise" = 0.005, *PCSK9* "Prime edits with scaffold" = 0.0057, *PCSK9* "Precision score" = 0.0104, *HBEGF* "Prime edits - precise" = 0.0123, *HBEGF* "Prime edits with scaffold" = 0.0018 *HBEGF* "Precision score" = 0.0027. **b** PRINS editing of mouse embryos using electroporation of PEn-pegRNA or PEn-rSp-pegRNA RNP complexes to install a 5 bp insertion into the endogenous *Map3k15* locus. Each data point represents a single embryo edited with either pegRNA ($n = 25$) or rSp-pegRNA ($n = 24$). Single embryos were analyzed by amplicon-seq and CRISPResso2 after 5 days of in vitro cultivation. "Prime edits with scaffold" was calculated as total number of amplicon-seq prime edited reads with scaffold integration per total prime edited reads. "Precision score" was calculated as total number of amplicon-seq reads with precise prime edit per overall editing. "Reads with precise edit" corresponds to a total number of precisely edited amplicon-seq reads in each embryo. Error bars represent mean ± SD. Statistical significance was determined using Student's t test (unpaired, two-tailed). $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Calculated P values: "Prime edits with scaffold" <0.0001, "Precision score" = 0.011, "Reads with precise edit" = 0.0319. Source data are provided as a Source Data file.

contrasted with that of the standard M-MLV RT used in regular PEs, where scaffold readthrough beyond the last two scaffold nucleotides is rare (Fig. 1e and Supplementary Fig. 4). After applying C96Me-pegRNA, this readthrough was decreased to 0.8% and 0.4% of total reads for PEn6d and PEn**, respectively (Fig. 5c), which corresponded to 1.7% and 0.7% of the total prime edited reads, for PEn6d and PEn**, respectively (Fig. 5d). These data show that RT readthrough with processive polymerases is frequent and can be efficiently blocked by 2′-O-Me. Lastly, we compared prime editing outcomes of C96Me-pegRNAs and unmodified pegRNAs using the nickase-based PE6d and PE**. As expected, the frequency of scaffold readthrough was much lower with nickase-based PEs. Nevertheless, the scaffold incorporation with unmodified pegRNAs was detected in ~0.8% of total sequenced reads for both PE6d and PE** (Fig. 5e), corresponding to 3.8% and 3.1% of prime edits for PE6d and PE**, respectively (Fig. 5f). The most abundant PE6d scaffold incorporation variant was a single substitution matching the third nucleotide of pegRNA scaffold (Supplementary Fig. 6b). After applying C96Me-pegRNA, the scaffold incorporation was virtually eliminated, reaching ~0.17% of total reads for PE6d and below the 0.1% detection limit of amplicon-seq for PE** (Fig. 5e). Scaffold insertion profiles for PE6d and PE** with pegRNA and C96Me-pegRNA are summarized in Supplementary Fig. 7. Together, our data demonstrate that the use of rSp-pegRNAs and C96Me-pegRNAs efficiently mitigates scaffold incorporation by nickase-based PEs, including highly processive variants.

Discussion

Scaffold incorporation is a by-product of prime editing that leads to introduction of undesired mutations downstream of the intended prime edit and thus compromises prime editing precision and safety. While the reported frequencies of scaffold incorporation are typically low, it is a common by-product of prime editing observed in numerous systems including cell lines, primary cells and zebrafish embryos[1–8]. Moreover, it is highly likely that scaffold incorporation will become more problematic as new and more processive prime editing systems emerge. This is exemplified by two recent studies where more processive RT variants allowed more efficient prime editing, but also led to increased scaffold incorporations[6,9]. Prior to this work, a universal strategy to mitigate scaffold incorporations was lacking. Previously, recoding of the scaffold directly upstream of the RTT to match the target locus sequence mitigated this issue[2,8]. However, such target-specific adjustments of the pegRNA scaffold sequence may impair its interaction with Cas9. Furthermore, while recent structural insights into PE suggest that only the last two to three nucleotides of the pegRNA scaffold are accessible to RT, reverse transcription may extend further into the scaffold. In such instances, mere recoding of the last few nucleotides of the scaffold does not suffice to prevent scaffold incorporations[8]. Both abasic spacers and 2′-O-Me precisely block RT directly at the end of RTT, and our strategy works irrespective of the target locus sequence and without the need to change the sequence of the pegRNA scaffold. Our extensive benchmarking of the modified pegRNAs demonstrates their robustness in multiple cell types, including primary human hepatocytes, where we achieved efficient and precise insertions with nuclease-based PRINS prime editing. Importantly, we show efficient blocking of RT, including its highly processive variants, with a single 2′-O-Me positioned at the last nucleotide of the pegRNA scaffold.

Both abasic and 2′-O-Me modifications offer unique advantages and drawbacks. The primary benefit of using abasic pegRNAs is the complete elimination of scaffold integration. In contrast, 2′-O-Me modification allows for some degree of readthrough, which is particularly noticeable with PRINS, although minimal when pegRNA with a longer template was used. Abasic pegRNAs tend to be less cost-effective compared to C96Me-pegRNAs, which incur no additional cost. Additionally, in the context of pegRNA synthesis, which is often challenging due to the large size of pegRNAs, rSp/C3 spacers are not typically standard catalog items and may require custom synthesis, resulting in higher costs and longer delivery times. Despite extensive testing across multiple cell types, we have not encountered any quality issues with abasic pegRNAs, however, the behavior and stability of gRNAs with abasic sites have not been studied in vivo. Therefore, we recommend using abasic pegRNA for pronounced cases of unwanted scaffold integrations, such as PRINS, where maximum stringency is required, and 2′-O-Me might not fully prevent scaffold readthrough.

Conversely, the cost and synthesis complexity of C96Me-pegRNAs are comparable to standard pegRNAs. Additionally, 2′-O-Me in gRNAs acts as a stabilizer and has been well-characterized, including in vivo and in therapeutic contexts[18,20]. Notably, highly modified gRNAs with multiple internal methylations exhibit increased activity in vivo, likely due to enhanced stability[20]. Thus, we recommend using C96Me-pegRNAs for nickase-based prime editing, where scaffold readthrough is generally lower and can be largely mitigated by 2′-O-Me, and for therapeutic applications of prime editing due to the well-understood behavior of 2′-O-Me in vivo.

Methods

Ethical statement

AstraZeneca has a governance framework and processes in place to ensure that commercial sources have appropriate patient consent and ethical approval in place for collection of the samples for research purposes including use by for-profit companies.

DNA constructs, synthetic pegRNAs and mRNAs

PE6d[6] and PE**[9] and their WT-Cas9 (H840) variants PEn6d and PEn** were generated by gene synthesis (GenScript). Synthetic pegRNAs without internal modifications were synthesized by IDT. Synthetic

https://doi.org/10.1038/s41467-025-58653-1

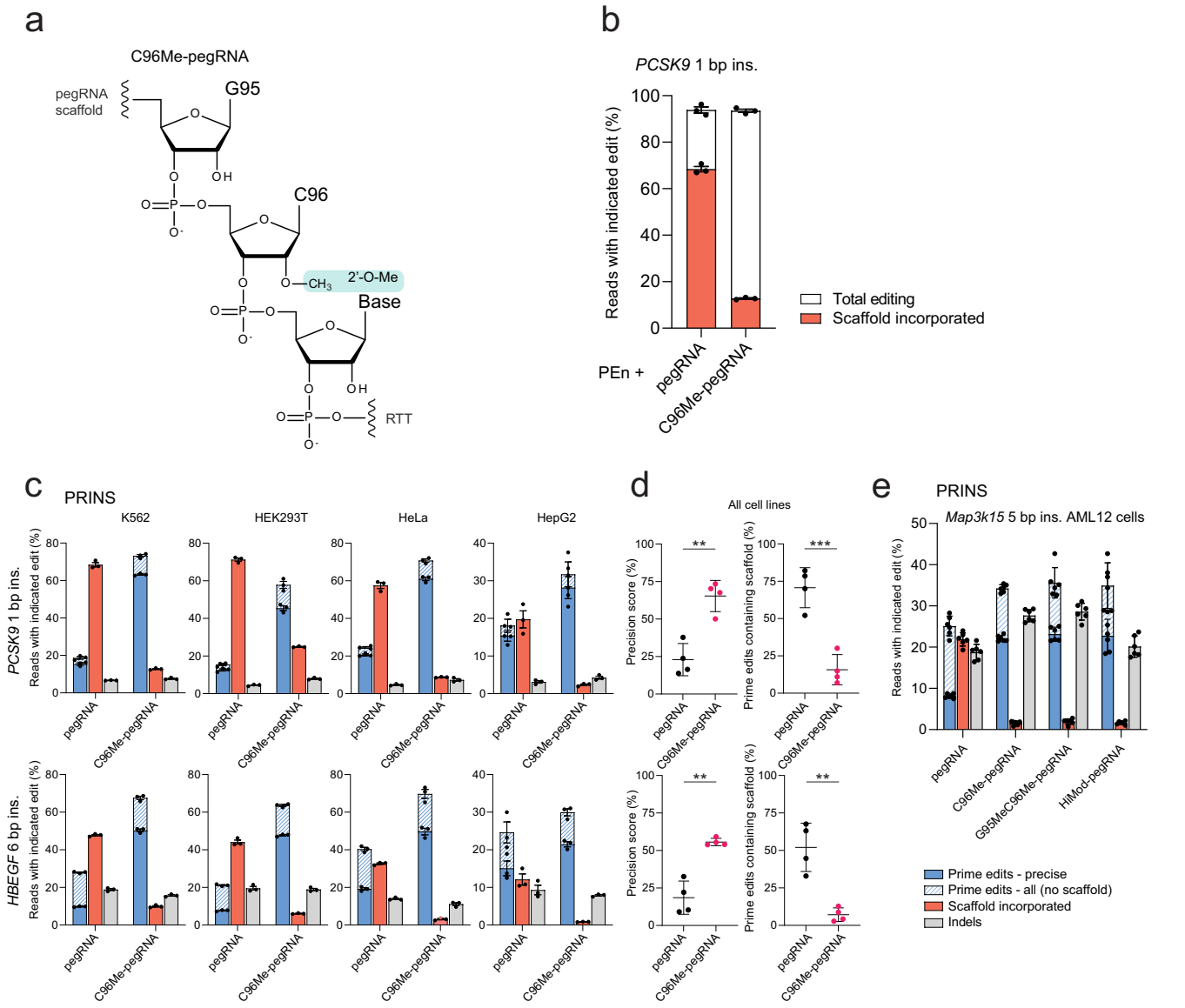


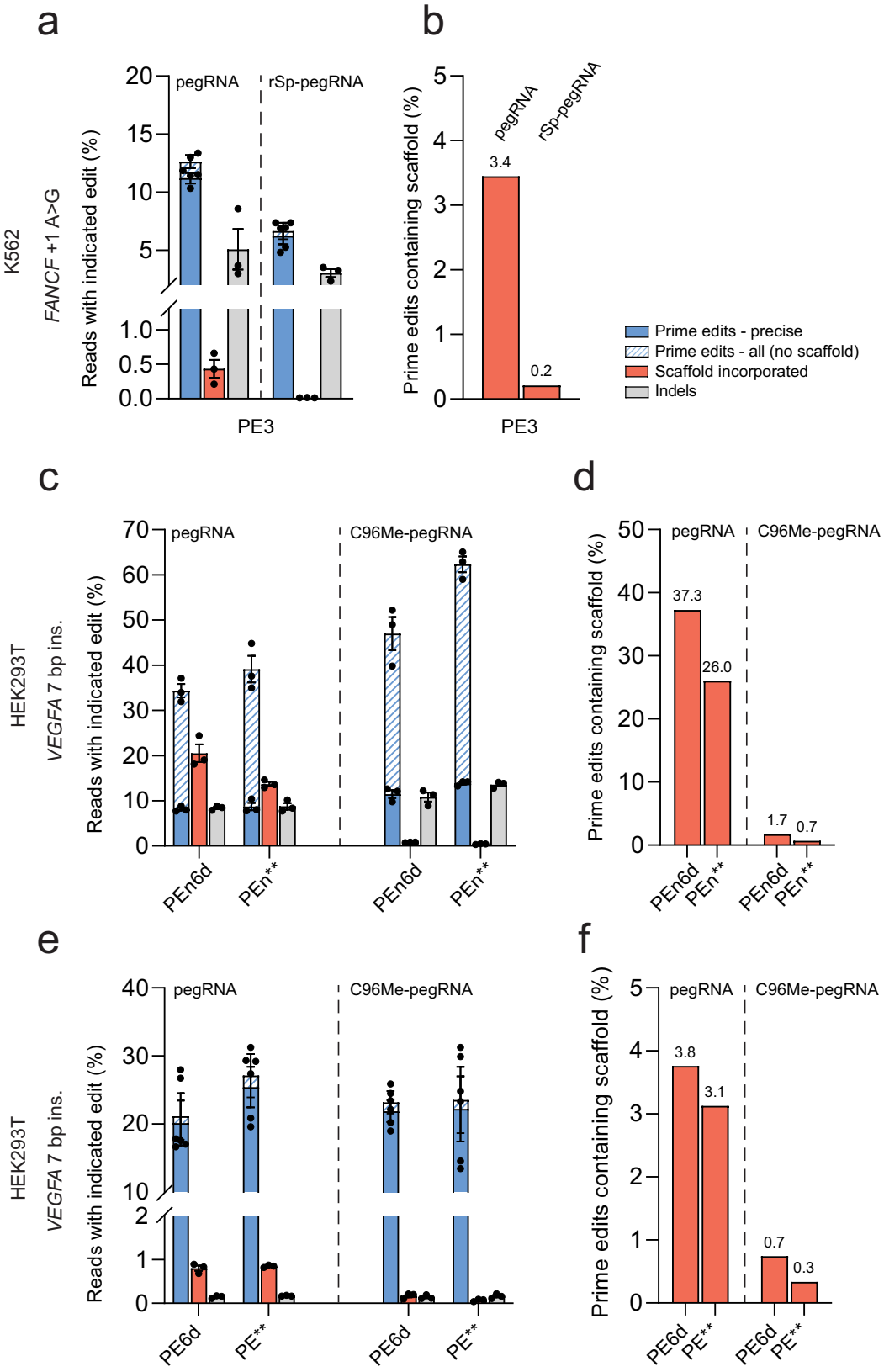
Fig. 4 | 2′-O-methylation of C96 in the pegRNA scaffold precisely blocks reverse transcription by M-MLV RT. a Chemical structure of C96 methylation within the scaffold of C96Me-pegRNA. C96 = cytosine at the position 96 of the SpCas9 pegRNA scaffold. G95 = guanine at the position 95 of the SpCas9 pegRNA scaffold. RTT = reverse transcription template. **b** Scaffold incorporation analysis by amplicon-seq of *PCSK9* PRINS editing in K562 cells electroporated with PEn mRNA and unmodified pegRNA or C96Me-pegRNA. **c** PRINS editing at *PCSK9* and *HBEGF* genomic loci in four different human cell lines (K562, HEK293T, HeLa and HepG2) using electroporation of PEn mRNA and synthetic pegRNA or C96Me-pegRNA to install indicated small insertions (ins.). Editing outcomes were analyzed by amplicon-seq and quantified using CRISPResso2 in the prime editing mode. Plots show mean ± SD of $n = 3$ biological replicates. **d** Quantification of data from (**c**). "Precision score" was calculated as total number of amplicon-seq reads with precise prime edit per overall editing. "Prime edits containing scaffold" was calculated as total number of amplicon-seq prime edited reads with scaffold integration per total prime edited reads. Error bars represent mean ± SD. Datapoints represent 4 different cell line experiments in (**c**). Statistical difference was determined using Student's *t* test (paired, two-tailed). Prime edits precise = intended prime edits; Prime edits all (no scaffold) = precise prime edits + prime edits co-occurring with indels not derived from scaffold; Scaffold incorporated = prime edits with at least one additional nucleotide derived from scaffold; Indels = non-prime edited insertions or deletions. $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$. Calculated *P* values: *PCSK9* precision score = 0.0016; *PCSK9* prime edits containing scaffold = 0.0009; *HBEGF* precision score = 0.0052; *HBEGF* prime edits containing scaffold = 0.0054. **e** PRINS editing of *Map3k15* locus in AML12 cells using lipofection of PEn mRNA and synthetic pegRNAs to install 5 bp insertion. Editing outcomes were analyzed by amplicon-seq and quantified using CRISPResso2 in the prime editing mode. Plots show mean ± SD of $n = 6$ biological replicates. Source data are provided as a Source Data file.

pegRNAs with abasic spacers were synthesized by Horizon Discovery or GenScript. Synthetic pegRNAs with internal 2′-O-Me were synthesized by IDT. Highly methylated *Map3k15* pegRNA was purchased from Axolabs. All pegRNA sequences and modifications used in this study are listed in Supplementary Data 1. Capped PEnmax mRNA was generated following T7 directed in vitro transcription using a linearized PEnmax DNA template. The in vitro transcription reaction produced fully modified mRNA, replacing uridines with N1-Methyl-pseudouridine and the mRNA was capped using TriLink's CleanCap-AG Cap1 analogue. The mRNA was subsequently column purified using MEGAClear transcription clean-up kit (ThermoFisher) and mRNA purity analyzed using a fragment analyzer (Agilent). The sequence of PEn corresponds to PEmax[2] with the H840A mutation reversed to the original histidine, restoring nuclease activity.

Protein expression and purification

The sequence of PEn with a C-terminal His-tag was cloned into pET24a. The expression plasmid was then transformed into *Escherichia coli*

https://doi.org/10.1038/s41467-025-58653-1



BL21lDE3 Star (Thermofisher) for use in protein production. Auto-induction protocol[24] was used for the over-production of PEn. Essentially, the culture was first grown over-night at 37 °C, before inoculation with 800 mL of ZYP autoinduction media, which was then grown at 37 °C with shaking until OD_{600} reached about 1–2. The temperature was then lowered to 18 °C and the culture was grown for a further 24 h, before harvesting the cells by centrifugation. Cell pellets were stored at −80 °C until further use. The cell pellets were then resuspended in 20 mM HEPES, pH 7.5, 500 mM NaCl, 1 mM DTT, 10% glycerol, and lysed by one pass through an Emulsiflex C3 (Avestin). The lysate was clarified by centrifugation at 20,000 × *g* for 20 min. The supernatant was supplemented with 10 mM imidazole and the lysate

Fig. 5 | Modified pegRNAs mitigate scaffold incorporation by PE3 and highly processive PE systems. a Editing of *FANCF* with PE3 in K562 cells. Cells were electroporated with RNPs and either pegRNA or rSp-pegRNA to install indicated substitution at the *FANCF* genomic locus. **b** Quantification of data from (**a**). "Prime edits containing scaffold" was calculated as total number of amplicon-seq prime edited reads with scaffold integration per total prime edited reads. **c** Editing with prime editing nucleases PEn6d and PEn** in HEK293T cells. Cells were transfected with plasmids expressing PEn6d or PEn** in combination with synthetic pegRNA or C96Me-pegRNA to install a 7 bp insertion (ins.) at the *VEGFA* genomic locus. **d** Quantification of data from (**c**). "Prime edits containing scaffold" was calculated as total number of amplicon-seq prime edited reads with scaffold integration per total prime edited reads. **e** Editing with prime editing nickases PE6d and PE** in HEK293T cells. Cells were transfected with plasmids expressing PE6d or PE** in combination with synthetic pegRNA or C96Me-pegRNA to install a 7 bp insertion (ins.) at the *VEGFA* genomic locus. Editing outcomes were analyzed by amplicon-seq and quantified using CRISPResso2 in the prime editing mode. **f** Quantification of data from (**e**). "Prime edits containing scaffold" was calculated as total number of amplicon-seq prime edited reads with scaffold integration per total prime edited reads. Editing outcomes were analyzed by amplicon-seq and quantified using CRISPResso2 in the prime editing mode. Plots show mean ± SD of $n = 3$ biological replicates. Prime edits precise = intended prime edits; Prime edits all = precise prime edits + prime edits co-occurring with indels; Scaffold incorporated = prime edits with at least one additional nucleotide matching scaffold; Indels = non-prime edited insertions or deletions. Source data are provided as a Source Data file.

was loaded onto a 5 mL HiTrap column (Cytiva) equilibrated in the same buffer. The column was washed with 20 column volumes of 20 mM HEPES pH, 7.5, 500 mM NaCl, 1 mM DTT, 10% glycerol, 20 mM imidazole, before elution with 300 mM imidazole. The eluted protein was diluted to about 200 mM NaCl, before further purification on a 5 mL HiTrap Heparin SP column (Cytiva). Finally, the protein was further purified by size exclusion chromatography on a Superdex 200 (26/60) column (Cytiva) equilibrated in a buffer consisting of 20 mM HEPES, pH 7.5, 300 mM NaCl, 1 mM DTT, and 10% glycerol. The peak containing PEn protein was pooled, concentrated to 10 mg/mL, flash-frozen in liquid nitrogen and stored at −80 °C until required.

Cell culture

HEK293T (ATCC, CRL-3216) were maintained in DMEM-GlutaMAX with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin. HeLa cells (ATCC, CCL-2) were cultured in MEM-HEPES-GlutaMAX with 10% FBS and 1% penicillin/streptomycin. HepG2 (ATCC, HB-8065) were cultured in MEM-HEPES-GlutaMAX with 10% FBS, 1x non-essential amino acids, 1 mM sodium pyruvate and 1% penicillin/streptomycin. K562 cells (ATCC CRL-243) were maintained in RPMI-1640-GlutaMAX supplemented with 10% heat-inactivated FBS, 20 mM HEPES, 1 mM sodium pyruvate and 1% penicillin/streptomycin. All reagents were purchased from Gibco. Cells were maintained at 37 °C in a 5% CO_2 atmosphere and regularly tested for mycoplasma contamination. Cell line identity was authenticated through STR profiling (IDEXX BioAnalytics). Female primary human hepatocytes (LifeNet Health, LOT: 2018274-01) were thawed one day before transfection according to manufacturer's instructions and transferred to 50 mL of Human hepatocyte thawing medium (LifeNet Health). Cells were pelleted at 150 × *g* for 5 min; the pellet was gently resuspended in 25 mL of Human hepatocyte plating medium and supplement (LifeNet Health) and pelleted again at 100 × *g* for 5 min. Cells were resuspended in seeding medium, and 80,000 cells/well were plated in 100 µL of medium into 96-well plates (Corning BioCoat, Collagen I coated). Before transfection, seeding medium was exchanged for human hepatocyte culture medium and supplement (LifeNet Health). AML12 cells (ATCC CRL-2254) were cultured in DMEM/F12 + GlutaMAX medium (Gibco, REF 31331-028) supplemented with 10% FBS (Gibco, A5256701), 1X Insulin-Transferrin-Selenium (ITS-G) (Gibco, 41400045), and Dexamethasone (Merck D4902-25MG) at a final concentration 40 ng/ml.

Plasmid and synthetic gRNA transfection

One day prior to transfections 12,500 HEK293T cells were seeded per well of a 96-well plate. Cells were transfected with FuGENE HD (Promega) using a 6:1 Fugene-to-DNA ratio and 100 ng of total DNA per 96 well (100 ng of PE6d, PEn6d, PE**, or PEn** editor plasmid). Briefly, 1 µL of plasmid (100 ng/µL) was mixed with 0.6 µL of Fugene and 8.4 µL of Opti-MEM (Gibco), incubated at room temperature for 15 min and added to cells. One day post DNA transfection, cells were transfected with Lipofectamine RNAiMAX (Invitrogen) and 2 pmol of synthetic gRNAs. Briefly, 0.5 µL of RNAiMAX were mixed with 8.3 µL of Opti-MEM (Gibco) to form a working RNAiMAX solution. In parallel, 2 pmol of synthetic gRNA were mixed with 5 µL of Opti-MEM (Gibco) to form a working gRNA solution. 5 µL of working RNAiMAX solution were mixed with 5 µL of working gRNA solution, incubated at room temperature for 5 min and added to cells. Cells were harvested for targeted amplicon-sequencing analysis after 48 h.

RNA and RNP electroporation

HEK293T, HeLa and HepG2 cells were seeded at 30% confluency 48 h prior to transfection. K562 cells were resuspended at 3.5×10^5 cells/mL 24 h before transfection. The day of the electroporation cells were collected and washed twice with phosphate buffered saline (PBS) and resuspended in nucleofection buffer (SF/SE Cell Line 96-well Nucleofector™ Kit; Lonza) containing 2 µg of PEn mRNA and 100 pmol of synthetic pegRNA, at 1×10^4 cells/µL. HEK293T, HepG2 and K562 cells (4×10^4, 4×10^4 and 2×10^5 cells/condition, respectively) were electroporated using the SF Cell Line 96-well Nucleofector™ Kit (Lonza) and the CM-130, CN-114 or FF-120 pulse code (Nucleofector 4D; Lonza), respectively. HeLa cells (4×10^4 cells/condition) were electroporated using the SE Cell Line 96-well Nucleofector™ Kit (Lonza) and the EH-100 pulse code (Nucleofector 4D; Lonza). Post-electroporation, cells were incubated at room temperature for 10 min without any disturbance and were thence transferred into pre-warmed medium in 96-well plates and incubated at 37 °C with 5% CO_2. For RNP nucleofection of K562 cells, 54 pmol of PE2 protein were combined with 100 pmol of pegRNA or rSp-pegRNA and 66 pmol of nicking gRNA in buffer SF and incubated at room temperature for 10 min to allow RNP formation. Subsequently, 5 µL of RNP complexes were added to 4×10^5 of cells in 15 µL of buffer SF. The final 20 µL suspension was electroporated using SF Cell Line 96-well Nucleofector™ Kit in Nucleofector 4D (Lonza) and pulse code FF-120. After nucleofection, cells were left undisturbed for 10 min and plated in 200 µL of pre-warmed growth media in a 96 well plate. Cells were collected for DNA extraction after 72 h.

RNA transfection of primary human hepatocytes and AML12 cells

Primary human hepatocytes were transfected 1 day after thawing using Lipofectamine MessengerMAX Transfection Reagent (Invitrogen) using the following protocol: Lipofectamine solution was prepared by mixing 0.3 µL of Lipofectamine MessengerMAX with 4.7 µL of Opti-MEM (Gibco) and incubated at room temperature for 10 min. Meanwhile, RNA solution was prepared by diluting 170 ng of PEn mRNA and 60 ng of pegRNA with Opti-MEM to a final volume of 5 µL. Both solutions were combined, incubated at room temperature for 15 min and added to cells. Cells were harvested for targeted amplicon-sequencing analysis after 48 h. One day prior transfection 5000 of AML12 cells were seeded per well of a 96-well plate, in 100 µL of DMEM/F12 + GlutaMAX medium (Gibco, REF 31331-028) supplemented with 10% FBS (Gibco, A5256701), 1X Insulin-Transferrin-Selenium (ITS -G) (Gibco, 41400045), and Dexamethasone (Merck D4902-25MG) at a final concentration 40 ng/m. Cells were

transfected using the following protocol: Lipofectamine solution was prepared by mixing 0.3 µL of Lipofectamine MessengerMAX (LMRNA008, Invitrogen™) with 4.7 µL of Opti-MEM I + GlutaMAX-I (Gibco, 51985-026) and incubated at room temperature for 10 min. Meanwhile, RNA solution was prepared by diluting 170 ng of PEn mRNA and 2 pmol of pegRNA with Opti-MEM to a final volume of 5 µL. Both solutions were combined, incubated at room temperature for 15 min and added to cells.

Mouse embryo electroporation

RNPs were formed at room temperature for 10 min in 10 µL of 2x Cas9 buffer (200 mM KCl, 40 mM HEPES) containing 12 µM PEn protein and 13 µM sgRNA. The final electroporation mix (20 µL) was prepared by combining 10 µL RNP mix with 10 µL of Opti-MEM (Gibco) and 50 pooled zona-intact zygotes from C57Bl6/N (Janvier labs). Electroporation was performed using BIORAD Gene Pulser Xcell Electroporation system with the following parameters: square wave protocol, voltage: 30 V, pulse length: 3 ms, number of pulses: 10, pulse interval 100 ms, cuvette length: 1 mm. The embryos were cultivated in 24-well cell culture plates for 5 days. For lysis, single embryos were placed in 200 µL PCR strips and lysed using 10 µL 1X Modified Gitschier buffer (0.2 M Tris pH 8.8, 0.1 M $(NH4)_2SO_4$, 50 mM $MgCl_2$, 1.7 µM SDS, and 0.1 mg/mL Proteinase K) at 37 °C for 1 h followed by 10 min heat inactivation at 85 °C. Four µL of lysates were used as PCR template to perform amplicon sequencing.

Genomic DNA extraction and amplicon sequencing

Cells were harvested using Quick Extract solution (Lucigen). Amplicons were generated with Phusion Flash High-Fidelity 2x Mastermix (F548, Thermo Scientific) or Q5 Hot Start High-Fidelity 2x Mastermix (M0492, NEB) in a 15 µL reaction, containing 1.5–2 µL of genomic DNA extract and 0.2 µM of target-specific primers with barcodes and adapters for next generation sequencing (NGS). All primer sequences are listed in Supplementary Data 2. PCR cycling conditions for Phusion Flash High-Fidelity 2x Mastermix were: 98 °C for 3 min, followed by 30 cycles of 98 °C for 10 seconds, 60 °C for 20 s, and 72 °C for 30 s. For Q5 Hot Start High-Fidelity 2x Mastermix the following PCR protocol was applied: 98 °C for 30 s, followed by 30 cycles of 98 °C for 10 s, 60 °C for 20 s, and 72 °C for 30 s, and final elongation at 72 °C for 2 min. Amplicons from mouse embryo lysates were generated using Q5 Hot Start High-Fidelity 2x Mastermix (M0492, NEB) in a 25 µL reaction, containing 4 µL of genomic DNA extract and 0.2 µM of target-specific primers with barcodes and NGS adapters. PCR program to generate amplicons from mouse embryos: 98 °C for 30 s, followed by 33 cycles of 98 °C for 10 s, 64 °C for 15 s, and 72 °C for 15 s, and final elongation at 72 °C for 2 min. All amplicons were purified using HighPrep PCR Clean-up System (MagBio Genomics). The size, purity, and concentration of amplicons were determined using a fragment analyzer (Agilent). To add Illumina indexes to the amplicons, samples were subjected to a second round of PCR. Indexing PCR was performed using KAPA HiFi HotStart Ready Mix (Roche), 0.067 ng of PCR template and 0.5 µM of indexed primers in the total reaction volume of 25 µL. PCR cycling conditions were 72 °C for 3 min, 98 °C for 30 seconds, followed by 10 cycles of 98 °C for 10 s, 63 °C for 30 s, and 72 °C for 3 min, with a final extension at 72 °C for 5 min. Samples were purified with the HighPrep PCR Clean-up System (MagBio Genomics) and analyzed using a fragment analyzer (Agilent). Samples were quantified using a Qubit 4 Fluorometer (Life Technologies) and subjected to sequencing using Illumina NextSeq system according to the manufacturer's instructions.

Bioinformatic analysis

Demultiplexing of the targeted amplicon sequencing data was performed using bcl2fastq software (Illumina). The fastq files were analyzed using CRISPResso2 V2.1.1 in the prime editing mode[25]. Detailed parameters are listed in the Supplementary Data 3.

Statistics & reproducibility

Data visualization and statistical analysis were conducted using GraphPad Prism 9 (GraphPad Software, Inc.) or JMP 14.1.0 (SAS Institute Inc.). Figure legends contain information on statistical tests, sample sizes, and *P* values. No data were excluded from the analyses. No statistical method was used to predetermine sample size. The experiments were not randomized. The investigators were not blinded during experiments and outcome assessment. Detailed statistical analyses are listed in the Supplementary Data 4.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All details and data to support the findings of this study are part of the manuscript or available in a publicly accessible repository. Amplicon sequencing samples are described in Supplementary Data 5 and source sequencing data are available in the NCBI Sequence Read Archive database, BioProject accession code PRJNA1226148. Source data are provided with this paper.

References

1. Anzalone, A. V. et al. Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019).
2. Chen, P. J. et al. Enhanced prime editing systems by manipulating cellular determinants of editing outcomes. *Cell* **184**, 5635–5652.e5629 (2021).
3. Petri, K. et al. CRISPR prime editing with ribonucleoprotein complexes in zebrafish and primary human cells. *Nat Biotechnol* https://doi.org/10.1038/s41587-021-00901-y (2021).
4. Levesque, S. et al. Marker-free co-selection for successive rounds of prime editing in human cells. *Nature Communications* **13**, 5909 (2022).
5. Peterka, M. et al. Harnessing DSB repair to promote efficient homology-dependent and -independent prime editing. *Nat Commun* **13**, 1240 (2022).
6. Doman, J. L. et al. Phage-assisted evolution and protein engineering yield compact, efficient prime editors. *Cell* **186**, 3983–4002.e3926 (2023).
7. Fiumara, M. et al. Genotoxic effects of base and prime editing in human hematopoietic stem cells. *Nature Biotechnology* https://doi.org/10.1038/s41587-023-01915-4 (2023).
8. Shuto, Y. et al. Structural basis for pegRNA-guided reverse transcription by a prime editor. *Nature* https://doi.org/10.1038/s41586-024-07497-8 (2024).
9. Ponnienselvan, K. et al. Addressing the dNTP bottleneck restricting prime editing activity. *bioRxiv*, 2023.2010.2021.563443 https://doi.org/10.1101/2023.10.21.563443 (2023).
10. Adikusuma, F. et al. Optimized nickase- and nuclease-based prime editing in human and mouse cells. *Nucleic Acids Res.* **49**, 10785–10795 (2021).
11. Jiang, T., Zhang, X.-O., Weng, Z. & Xue, W. Deletion and replacement of long genomic sequences using prime editing. *Nature Biotechnology* https://doi.org/10.1038/s41587-021-01026-y (2021).
12. Tao, R. et al. WT-PE: Prime editing with nuclease wild-type Cas9 enables versatile large-scale genome editing. *Signal Transduct Target Ther.* **7**, 108 (2022).
13. Kweon, J. et al. Targeted genomic translocations and inversions generated using a paired prime editing strategy. *Mol. Ther.* **31**, 249–259 (2023).
14. Li, X. et al. Development of a versatile nuclease prime editor with upgraded precision. *Nat Commun* **14**, 305 (2023).
15. Takeshita, M., Chang, C. N., Johnson, F., Will, S. & Grollman, A. P. Oligodeoxynucleotides containing synthetic abasic sites. Model

substrates for DNA polymerases and apurinic/apyrimidinic endonucleases. *Journal of Biological Chemistry* **262**, 10171–10179 (1987).
16. Singh, P., Schimenti, J. C. & Bolcun-Filas, E. A mouse geneticist's practical guide to CRISPR applications. *Genetics* **199**, 1–15 (2015).
17. Nag, A. et al. Human genetics uncovers *MAP3K15* as an obesity-independent therapeutic target for diabetes. *Science Advances* **8**, eadd5430 (2022).
18. Hendel, A. et al. Chemically modified guide RNAs enhance CRISPR-Cas genome editing in human primary cells. *Nature Biotechnology* **33**, 985–989 (2015).
19. Maden, B. E., Corbett, M. E., Heeney, P. A., Pugh, K. & Ajuh, P. M. Classical and novel approaches to the detection and localization of the numerous modified nucleotides in eukaryotic ribosomal RNA. *Biochimie* **77**, 22–29 (1995).
20. Finn, J. D. et al. A Single Administration of CRISPR/Cas9 Lipid Nanoparticles Achieves Robust and Persistent In Vivo Genome Editing. *Cell Rep*. **22**, 2227–2235 (2018).
21. Liang, S. Q. et al. Genome-wide profiling of prime editor off-target sites in vitro and in vivo using PE-tag. *Nat Methods* **20**, 898–907 (2023).
22. Liu, P. et al. Increasing intracellular dNTP levels improves prime editing efficiency. *Nature Biotechnology* https://doi.org/10.1038/s41587-024-02405-x (2024).
23. Levesque, S., Cosentino, A., Verma, A., Genovese, P. & Bauer, D. E. Enhancing prime editing in hematopoietic stem and progenitor cells by modulating nucleotide metabolism. *Nature Biotechnology* https://doi.org/10.1038/s41587-024-02266-4 (2024).
24. Studier, F. W. Protein production by auto-induction in high-density shaking cultures. *Protein Expression and Purification* **41**, 207–234 (2005).
25. Clement, K. et al. CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nature Biotechnology* **37**, 224–226 (2019).

Acknowledgements

We thank the AstraZeneca Discovery Sciences Genome Engineering team for support and input on this work. We thank Steve Rees and Mike Snowden for supporting this project. We are grateful to the AstraZeneca NGS & Transcriptomic team for support with targeted amplicon sequencing. P.A. is funded by the European Union's Horizon Europe research and innovation programme under grant agreement No 101057659 (EDITSCD). L.D. and A.S. are funded by Promega Corporation.

Author contributions

P.A., L.D., N.A., M.M. and M.P. conceptualized the study and prepared the manuscript. P.A., L.D., N.A. and M.P. performed most of the experimental work and analysis with support from H.M., K.M.B., A.L.L., A.S., E.G., S.M., G.T., P.P.H., S.Š. and M.F.

Competing interests

P.A., L.D., H.M., K.M.B., A.L.L., A.S., E.G., S.M., G.T., P.P.H., S.Š., M.F., N.A., M.M., M.P are employees of AstraZeneca and may be AstraZeneca shareholders. L.D. and A.S. are funded by Promega corporation. AstraZeneca filed patents related to this work (WO2021204877A2 and WO2023052508A2).

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-58653-1.

Correspondence and requests for materials should be addressed to Marcello Maresca or Martin Peterka.

Peer review information *Nature Communications* thanks Shaohua Yao and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.