



OPEN A buffer-tuning strategy to profile domain-specific activity of chimeric I-TevI/CRISPR gene editors *in vitro*

Kurt W. Loedige¹, Alexa L. White¹, Thomas A. McMurrough², Brent E. Stead² & David R. Edgell¹✉

Protein-DNA interactions can be manipulated *in vitro* by changing buffer conditions. Here, we develop a methodology to map the cleavage preferences of chimeric gene editors that are fusions of the I-TevI nuclease domain to CRISPR nucleases by manipulating *in vitro* salt concentrations. We found that DNA cleavage by the I-TevI (Tev) nuclease domain at CNNNG sites was de-coupled from the gRNA-targeted site in low salt buffers. For TevCas12a, this non-targeted cleavage activity was enriched at Tev CNNNG cleavage motifs optimally positioned within a 30-bp window upstream of a Cas12a TTTV PAM site. Non-targeted cleavage did not require Cas12a nuclease activity or specific Cas12a gRNA targeting. Similar non-targeted products were observed in low salt buffer conditions for TevSaCas9, Tev-meganuclease and Tev-zinc finger editors. Cas12a and SaCas9 activity at gRNA-directed sites and sites with multiple mismatches were also sensitive to buffer salt concentration. Oxford Nanopore sequencing revealed a remarkably similar Tev CNNNG cleavage preference at different salt concentrations and in different fusion contexts, emphasizing the robustness and specificity of Tev activity. More generally, our work highlights the sensitivity of gene editors to *in vitro* reaction conditions and how these conditions can be leveraged to functionally dissect the activity of individual domains of chimeric gene editors.

Genome editing nucleases are programmable, site-specific DNA endonucleases that enable targeted genome modification^{1,2}. A critical aspect of their application is understanding DNA binding and cleavage specificity, which informs the selection of on-target sites that support high activity while minimizing activity at mismatch-containing off-target sites. Profiling the specificity of gene editors often relies on *in vitro* methodologies that allow activity determination on a larger number substrates than is possible *in vivo*^{3–5}. However, *in vitro* protein-DNA interactions are sensitive to buffer conditions^{6–9}. In particular, modulating ionic buffer strength has been leveraged as a tool to probe the mechanisms and kinetics of gene editors, including zinc-finger nucleases¹⁰, TALE domains^{11,12}, Cas9^{13,14}, and Cas12a¹⁵.

We previously showed that the GIY-YIG nuclease domain and interdomain linker from the homing endonuclease I-TevI (Tev) could be fused to unrelated DNA-binding modules (zinc fingers, TALEs, meganucleases and Cas9 orthologs) to create chimeric gene editors^{16–20}. In these editors, the DNA-binding module dictates target site selection, while Tev cleaves at a CNNNG motif that is appropriately spaced from the target site, with some differences in spacing for each fusion context^{16–18}. In particular, the fusion of Tev to SpCas9 from *Streptococcus pyogenes*, a class II type II-B nuclease¹⁹, created an RNA-guided dual nuclease (TevSpCas9) that makes two double-strand breaks (DSBs): one at a Tev CNNNG motif upstream of the 5' end of the gRNA target site (producing a 2-nt 3' overhang) and another at the Cas9 cut site (yielding a blunt end). In mammalian cells, TevSpCas9 editing generates defined length deletions with high frequency and no detectable exonucleolytic processing of the DNA ends¹⁹.

We sought to expand on the dual RNA-guided nuclease architecture by fusing Tev to the N-terminus of *Acidaminococcus* sp. Cas12a (TevCas12a) (Fig. 1A). Cas12a (Cpf1) is a class II, type V-A CRISPR nuclease that cleaves double-strand DNA (dsDNA) at sites specified by a 20-nucleotide crRNA adjacent to a TTTV (V=A/C/G) protospacer-associated motif (PAM) sequence^{21,22}. Cas12a and related proteins have been adapted for gene editing in mammalian systems and for nucleic acid diagnostic applications^{23–26}. We anticipated that cleavage by TevCas12a would generate two DSBs: one by Tev cleavage at a CNNNG motif to create a 2-nt 3' overhang, and the other by Cas12a to create a 4-nt 5' overhang.

¹Department of Biochemistry, Schulich School of Medicine & Dentistry, Western University, London, ON N6A 5C1, Canada. ²Specific Biologics Inc., Toronto, ON M5G 0B7, Canada. ✉email: dedgell@uwo.ca

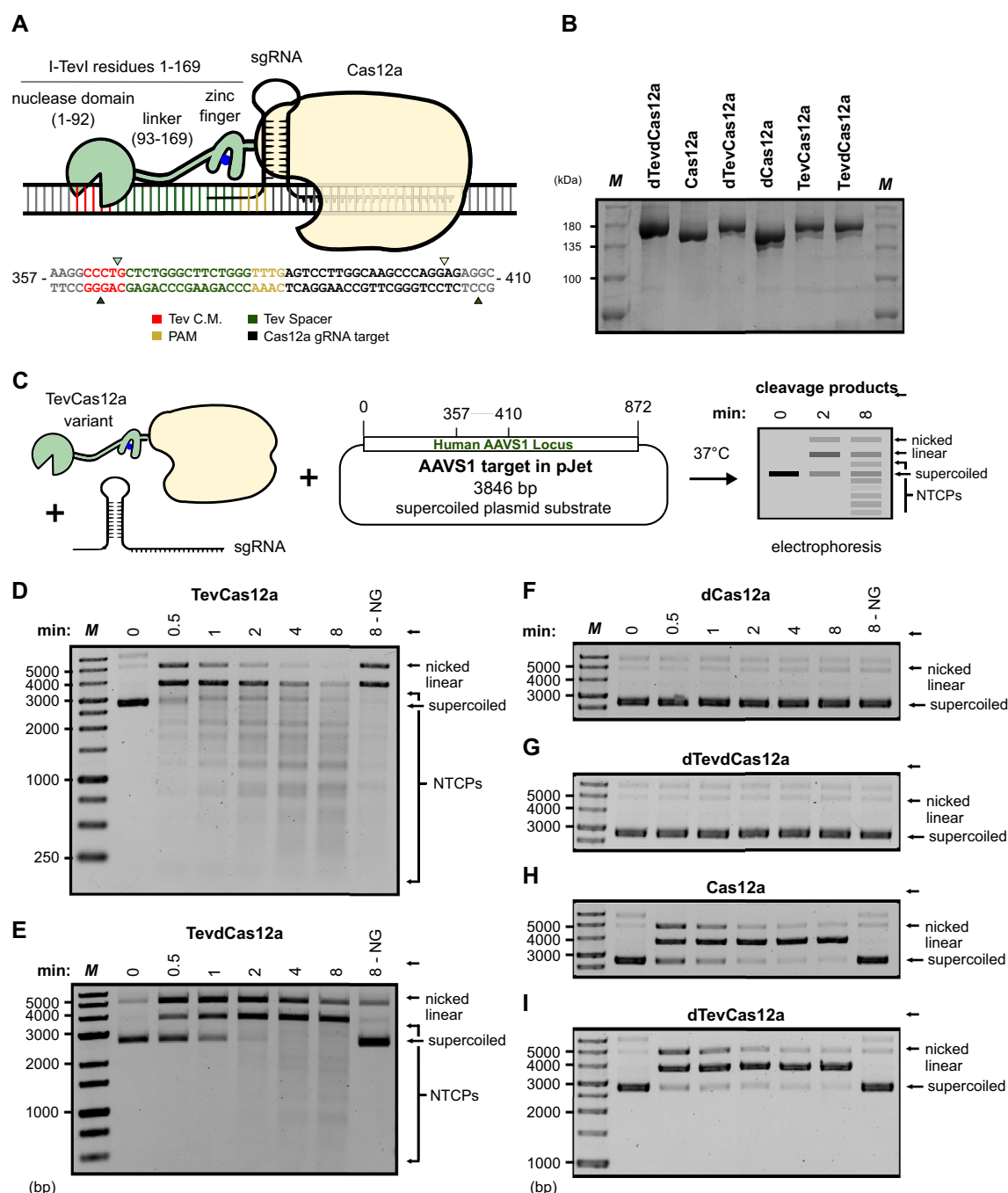


Fig. 1. Non-targeted cleavage of supercoiled DNA by TevCas12a. (A) Schematic diagram of TevCas12a with annotated components alongside the pJET-AAVS1 target site sequence. Triangles indicated Tev or Cas12a cleavage sites. (B) Purified TevCas12a and Cas12a protein variants. "d" represents catalytically inactive (dead) domains. (C) Schematic diagram of *in vitro* cleavage reactions: purified proteins are formed into RNPs by incubation with targeted or non-targeted gRNA, incubated with supercoiled plasmid at 37°C and products resolved by agarose gel electrophoresis. (D) Time-course cleavage of pJET-AAVS1 with TevCas12a/gRNA_{AAVS1} showing both targeted Cas12a cleavage products and non-targeted cleavage products (NTCPs). M, molecular weight ladder with sizes in bp indicated on left. NG, no guide. (E–I) Time-course cleavage assays with pJET-AAVS1 and (E) TevdCas12a, (F) dCas12a, (G) dTevdCas12a, (H) Cas12a, or (I) dTevCas12a complexed with gRNA_{AAVS1}. Labeled as in panel D. Equivalent time-course assays with a non-targeted gRNA (gRNA_{CFTR}) are provided in Supplementary Fig. S2A–S2B. Uncropped gel images for panels B–I are shown in Supplementary Fig. S7.

Here, we characterize TevCas12a activity *in vitro* and unexpectedly find that dsDNA cleavage activity is uncoupled from the gRNA-targeted site in specific, low salt buffer conditions. This non-targeted cleavage activity requires Tev activity and Cas12a DNA binding, but not Cas12a nuclease activity. Using target site knockouts and Oxford Nanopore sequencing, we show that TevCas12a cleaves substrates at Tev CNNNG motifs that are optimally spaced 20-bp from a Cas12a TTTV PAM site. Non-targeted cleavage products were also observed *in vitro* with TevSaCas9, Tev-zinc finger nucleases, and Tev-meganuclease fusions. Our results demonstrate how Tev and Cas12a use different strategies to search for target sites and how these activities can be leveraged to map cleavage preferences of chimeric gene editors.

Results

Non-targeted cleavage of dsDNA substrates by TevCas12a

To construct TevCas12a, we fused I-TevI residues 1–169 encompassing the GIY-YIG nuclease domain and interdomain linker to the N-terminus of *Acidaminococcus* sp. Cas12a (TevCas12a), as well as to its catalytically dead variant D908P (TevdCas12a) (Fig. 1A,B). During our initial characterization of TevCas12a, we noted that the cleavage activity was sensitive to *in vitro* buffer conditions. In particular, low cleavage activity was observed in reactions with greater than 100 mM sodium chloride. To further characterize TevCas12a activity, we created a supercoiled plasmid substrate (pJET-AAVS1) with a ~1kb insert of the AAVS1 safe harbour site that was PCR amplified from human genomic DNA (Fig. 1C). When pJET-AAVS1 was incubated with RNPs consisting of purified TevCas12a (Fig. 1B) and an *in vitro* transcribed gRNA_{AAVS1} in buffer containing 50 mM sodium chloride and 1 × rCutSmart buffer, we observed robust on-target activity as well as multiple cleavage products ranging in size from ~100–3000-bp (Fig. 1D). These products were inconsistent with the size of the predicted TevCas12a linear products based on the position of the gRNA_{AAVS1} target site in the AAVS1 sequence (3.801-kb and a small 0.045-kb product corresponding to the fragment between the Tev and Cas12a cleavage sites) (Fig. 1C). We termed these cleavage products non-targeted cleavage products (NTCPs). At the end of the 8 min timecourse at a TevCas12a:pJET-AAVS1 ratio of 20:1, >95% of pJET-AAVS1 was converted to NTCPs (Fig. 1D). Lowering the TevCas12a:pJET-AAVS1 ratio reduced the appearance of NTCPs (Fig. S1E–J).

Interestingly, we also observed cleavage of pJET-AAVS1 into a discrete linear product and NTCPs with TevCas12a in the absence of the gRNA_{AAVS1} (apo-TevCas12a) after 8 min, albeit at a lower rate than when reactions were performed with TevCas12a/gRNA_{AAVS1} (Fig. 1D, NG (no guide)). NTCPs were also observed when we targeted TevCas12a to isolated linear λ DNA with a gRNA targeting gene *N* (gRNA_{A λ}) (Fig. 2A) or to a lesser extent with apo-TevCas12a at high protein:DNA ratios (Fig. 2C). The non-targeted activity of TevCas12a was enhanced in reactions containing 1 × CutSmart where sodium chloride concentrations were 100 mM or less, whereas the on-target gRNA-directed cleavage fraction was highest in buffers with up to 200 mM sodium chloride concentrations (Fig. 2A and B). As rCutSmart (NEBuffer 4) contains 50 mM potassium acetate (K⁺), these sodium chloride concentrations represent additive increases in total monovalent cations (e.g., 100 mM Na⁺ corresponds to 150 mM total monovalent cations).

Cleavage of dsDNA by Cas12a proceeds by nicking of the non-targeted DNA strand followed by conformational rearrangement of the RuvC domain to nick the targeted DNA strand^{21,22}. With a supercoiled plasmid substrate, these steps are visible as a nicked intermediate and linear product, and can be estimated by first-order reaction rates k_1 and k_2 , respectively (Table 1, Fig. S1A–D). However, with an active TevCas12a fusion that generates multiple NTCPs, individual nicking reaction products are difficult to distinguish and thus the overall reaction rate (k_{obs}) is estimated from depletion of supercoiled plasmid. Notably, TevCas12a timecourses were performed at a 10:1 enzyme:substrate ratio, whereas Cas12a, TevdCas12a and dTevCas12a timecourses were performed at 20:1, so the reported k_{obs} likely underestimates the true TevCas12a reaction rate. Even under these conditions, the TevCas12a k_{obs} rate (0.112 ± 0.010 nM min^{−1}) is very similar to the Cas12a k_1 rate (0.118 ± 0.016 nM min^{−1}) (Table 1). To help distinguish NTCPs from Cas12a on-target cleavage, we used the TevdCas12a fusion (Figs. 1E, S1B) and found that the k_{obs} rate was reduced by ~1.8-fold (0.067 ± 0.009 nM min^{−1}). The difference in reaction rates for TevdCas12a versus TevCas12a could be related to the extremely slow dissociation rate of dCas12a from dsDNA in Mg²⁺-containing buffer²⁷. Profiling the temperature dependence revealed a reduction of 50% and 75% in the appearance of non-targeted products at 30° C and 18° C as compared to 37° C (Fig. S2C).

Non-targeted cleavage is independent of gRNA targeting

To test the dependence of NTCPs on specific gRNA-target DNA interactions, we designed a gRNA to target a site in the human cystic fibrosis transmembrane conductance regulator (CFTR). When TevCas12a/gRNA_{CFTR} or TevdCas12a/gRNA_{CFTR} was incubated with pJET-AAVS1, we observed NTCPs at approximately the same rate as with the gRNA_{AAVS1} (Fig. S2A–B, compare with Fig. 1D, E). Spiking gRNA_{AAVS1} into a reaction of apo-TevCas12a and λ DNA at 8 min stimulated the appearance of NTCPs (Fig. 2D, compare with Fig. 2C). This result shows that gRNA stimulates NTCP generation, but that the stimulatory effect is not dependent on targeted gRNA-DNA interactions. To further illustrate this effect, we quantified λ DNA conversion to NTCPs across the three conditions: apo-TevCas12a (Fig. 2C), apo-TevCas12a with gRNA_{AAVS1} spiked in at 8 min (Fig. 2D), and TevCas12a with a non-targeted gRNA_{AAVS1} included from the start (Fig. 4D). The resulting data (Fig. S3) confirm that both spiked and non-targeted gRNAs reproducibly stimulate NTCP formation relative to apo-TevCas12a. The decrease of free gRNA at later time points reflects the incorporation of gRNA into RNP complexes over the course of the reaction. Moreover, the appearance of multiple NTCPs cannot be explained by cleavage of TevCas12a/gRNA_{AAVS1/CFTR} at off-target sites. Previous work has shown that Cas12a can nick dsDNA with gRNA-target mismatches of 4 or less²⁸; there are no sites with 6 or fewer mismatches in pJET-AAVS1 to gRNA_{CFTR}, and no sites with 7 or fewer mismatches to gRNA_{AAVS1}.

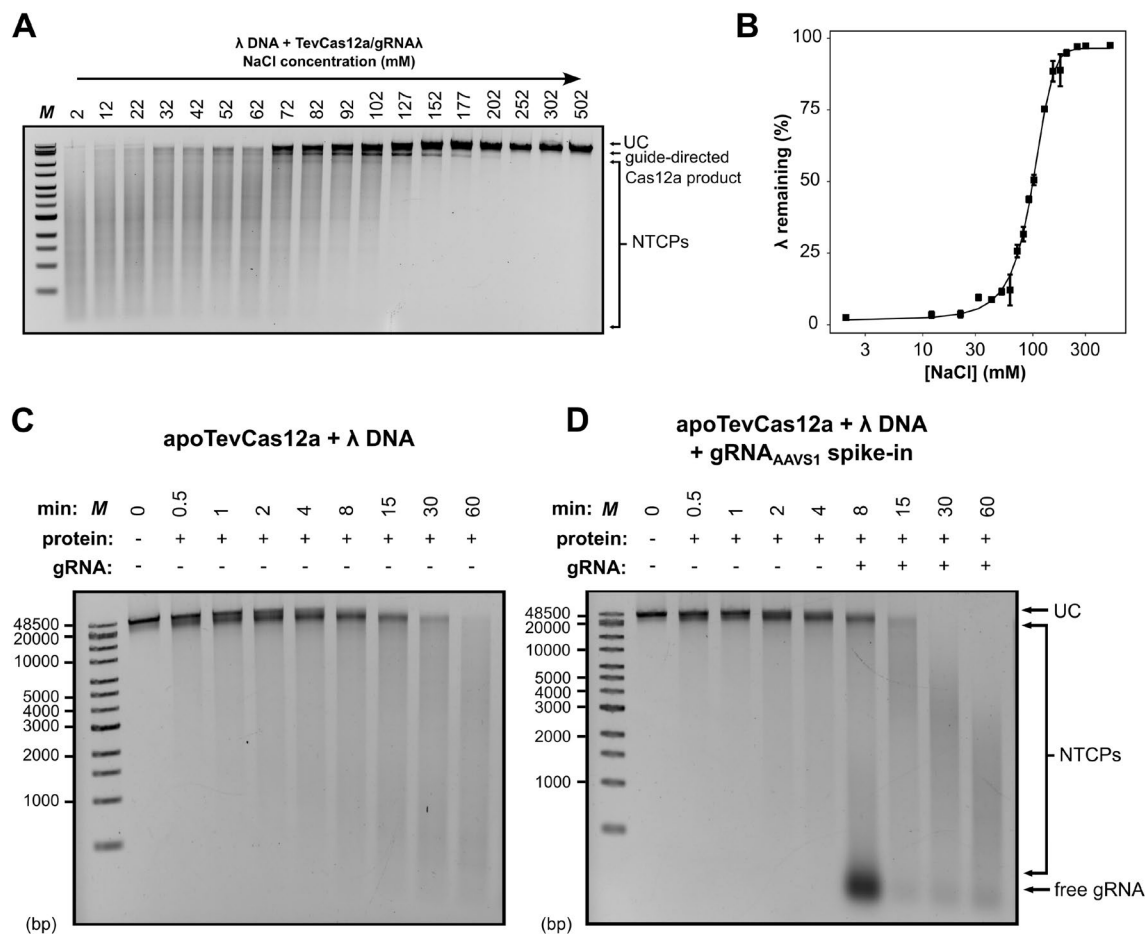


Fig. 2. Effect of NaCl concentration on CutSmart-buffered TevCas12a cleavage of λ DNA. **(A)** Representative agarose gel of titration of sodium chloride concentration in reactions with λ DNA and TevCas12a/gRNA_λ. The final sodium chloride concentration in each reaction is indicated above the indicated lane. Products and unreacted λ genome (UC) are indicated on the right of the gel. M, molecular weight marker. NTCPs, non-targeted cleavage products. **(B)** Plot of reaction sodium chloride concentration versus unreacted λ DNA remaining. Each point is mean of three independent replicates with error bars indicating standard deviation from the mean. The data were fit to a non-linear regression model. **(C)** Reaction of apoTevCas12a and λ DNA at the indicated time points. **(D)** Spike-in of gRNA to reaction of apoTevCas12a and λ DNA. Uncropped gel images for panels A–D are shown in Supplementary Fig. S8.

Enzyme	k_1	k_2	k_{obs}
Cas12a	0.118 ± 0.016	0.035 ± 0.007	N.D.
TevCas12a*	N.D.	N.D.	0.112 ± 0.010
TevdCas12a	N.D.	N.D.	0.067 ± 0.009
dTevCas12a	0.053 ± 0.011	0.038 ± 0.011	N.D.

Table 1. Observed reaction rates for targeted (k_1 and k_2) and non-targeted (k_{obs}) cleavage of pJET-AAVS1 (nM DNA/sec) for all proteins at 20:1, except TevCas12a at 10:1 (*) protein:DNA.

Non-targeted cleavage requires Tev but not Cas12a nuclease activity

We next generated an active site mutant in the Tev nuclease domain (R27A, dTev)^{29,30} to determine the dependency of NTCPs on Tev nuclease activity. With a TevdCas12a mutant (Fig. 1B, E), we found a reduction in the rate and extent of NTCPs (Fig. S1B). Nicked and linear products consistent with gRNA-targeted cleavage at the AAVS1 site by the Tev domain were still observed (Fig. 1E). No cleavage products were observed with a dCas12a mutant (Fig. 1B, F). However, with the reciprocal dTevCas12a mutant (Fig. 1B, I) that knocked out the Tev active site, NTCPs were abolished while on-target gRNA directed Cas12a cleavage was not impacted (Fig. 1I, Fig. S2D). The double mutant, dTevdCas12a (Fig. 1B), abolished both NTCPs and on-target gRNA cleavage (Fig. 1G). The appearance of NTCPs with the non-targeted CFTR gRNA was also dependent on an

active Tev nuclease domain and were reduced with the TevCas12a mutant, similar to that observed with gRNA_{AAVS1} (Fig. S2A–B).

Taken together, these data show that by simply changing *in vitro* salt concentration, TevCas12a activity can be modified to a non-targeted cleavage activity that does not require a programmed gRNA–DNA interaction. The non-targeted products and specific Cas12a–gRNA directed products appear at high protein:DNA ratios, consistent with single-turnover kinetics. The differences in salt sensitivity for the NTCPs and the gRNA-directed products suggest that these two activities result from different DNA cleavage events. Inactivation of Tev nuclease activity abolishes the appearance of NTCPs, whereas inactivation of Cas12a nuclease reduces but does not abolish the appearance of NTCPs. These observations rule out Cas12a collateral cleavage activity as a source of NTCPs. Moreover, the appearance of discrete products rather than complete plasmid degradation suggests that the non-targeted cleavage activity exhibits sequence specificity, consistent with the CNNNG preference of the Tev nuclease domain.

Tev CNNNG motifs are required for non-targeted cleavage

To further dissect the DNA target sequence requirements for the NTCPs, we next created a series of 500-bp substrates derived from pJET-AAVS1 where we mutated the Tev CNNNG cleavage motifs (Δ CNNNG) to non-cleavable sites by mutating the first C of the motif, the Cas12a TTTV PAM sites (Δ TTTV) by mutating the first T of the PAM, or both (Fig. 3A). In all cases, the on-target gRNA binding site, TTTV PAM and Tev CNNNG site (except for Δ CNNNG substrate) were retained. As shown in Fig. 3B with the wild-type substrate, the predicted on-target gRNA directed cleavage products and NTCPs were observed at 0.5 min, with complete substrate cleavage by 15 min. With the Δ CNNNG substrate that knocked out all CNNNGs, NTCPs were not observed (Fig. 3C). Interestingly, NTCPs were observed with the Δ TTTV substrate (Fig. 3D), although to a lesser extent and differently patterned than with the wild-type substrate (Fig. 3B). NTCPs and the Tev cleavage product 1 (TP1) were observed in the absence of an added gRNA, indicating that apo-TevCas12a retains DNA-binding activity at the protein:DNA concentrations used in the experiment (Fig. 3B, D). With the double Δ CNNNG + Δ TTTV knockouts, we observed only the on-target Tev and Cas12a cleavage products (Fig. 3E). Collectively, these data show that NTCPs require Tev CNNNG motifs and are enhanced with substrates that also contain Cas12a TTTV PAM sites.

Oxford Nanopore sequence mapping confirms CNNNG-dependence of non-targeted cleavage

To identify the NTCPs, we used Oxford Nanopore sequencing to map linear cleavage products from cleavage assays with λ DNA and TevCas12a/gRNA $_{\lambda}$ or with TevCas12a/gRNA_{AAVS1} (Fig. 4). In each case, NTCPs were visible at 0.5 min and complete cleavage of the λ genome in to NTCPs by 8 min (Figs. 4B and 5D). We also observed NTCPs in the absence of a gRNA, consistent with data from cleavage of AAVS1 linear substrates. In experiments with TevCas12a and (gRNA $_{\lambda}$), the on-target cleavage product mapped to the expected position (Fig. 4C). This product was not present in reactions with TevCas12a and gRNA_{AAVS1} (Fig. 4E). We sequenced reactions with Oxford Nanopore that were stopped after 2 minutes, and identified the Tev CNNNG cleavage motif from reads from experiments with the on-target gRNA $_{\lambda}$ and non-targeted gRNA_{AAVS1} (Fig. 4F). Changing the NaCl concentration in the reaction buffer did not alter the Tev cleavage preference on λ DNA, as a highly similar CNNNG motif preference was observed under both low (12 mM) and high (102 mM) NaCl conditions (Fig. S5A–B). Moreover, we also identified a T/A preference at position 7 (2 bases downstream of the G of the CNNNG motif) (Fig. 4F) that is consistent with base preference for the wild-type I-TevI enzyme and other chimeric Tev-based nucleases^{18,31}. Using CNNNG sequences present in the mapped cleavage products, we determined that the preference for the NNN central triplet was very similar to that determined from previous studies (Fig. 4G), where GC-rich NNN triplets are cleaved less efficiently than AT-rich triplets¹⁸.

Further analysis of the Oxford Nanopore sequencing and mapping data revealed a relationship between the frequency of cleavage at CNNNG motifs and the distance to the nearest TTTV PAM sequence (Fig. 5). These findings indicate a decreasing activity on Tev CNNNG motifs as the spacing increases upstream from the TTTV PAM. This data is consistent with studies showing that the I-TevI nuclease displays a distance preference for CNNNG motifs spaced appropriately from the DNA-binding domain³². A similar distance relationship was found for fusion of the I-TevI nuclease to zinc fingers, TALE domains, or meganucleases^{16–18}. We propose that scanning of DNA substrates for PAM sites by Cas12a and the subsequent steps for R-loop formation provide sufficient residence time for the Tev nuclease domain to cleave at CNNNG sites.

Cas12a and SaCas9 cleavage is sensitive to salt concentrations

It was recently shown that Cas12a searches for TTTV PAM sites by facilitated diffusion along a DNA substrate and that the rate of diffusion slows at lower salt concentrations^{15,33}. These observations suggest that Tev-directed NTCPs may result from Cas12a searching for TTTV PAM sites under the different salt concentrations used in the *in vitro* assays. To test this idea, we designed a 2000-bp linear “Cas-OFF” substrate that contained an overlapping, on-target AAVS1 Cas12a and SaCas9 site, as well as five sites that varied by one to three nucleotide transversion substitutions (Fig. 6A, S6A). The positions of the substitutions were PAM distal (>15 from the PAM), PAM medial (between 8–15 from the PAM), and PAM proximal (between 1–8 from the PAM) (Fig. 6A, S6A). Cleavage of each site by Cas12a would generate a discrete length product. When purified Cas12a/gRNA_{AAVS1} was incubated with the substrate across a 250-fold range of sodium chloride concentrations, we observed cleavage at each of the sites in the order on-target > PAM distal > PAM medial > PAM proximal (Fig. 6B–C). This cleavage preference is consistent with the tolerance of Cas12a/gRNA to target site mismatches^{34,35}.

We performed a similar experiment with SaCas9, titrating salt concentration in a reaction containing 1 $\times$ CutSmart and a substrate with on-target and mis-matched target sites (Fig. S6B–C). Similar to that observed with

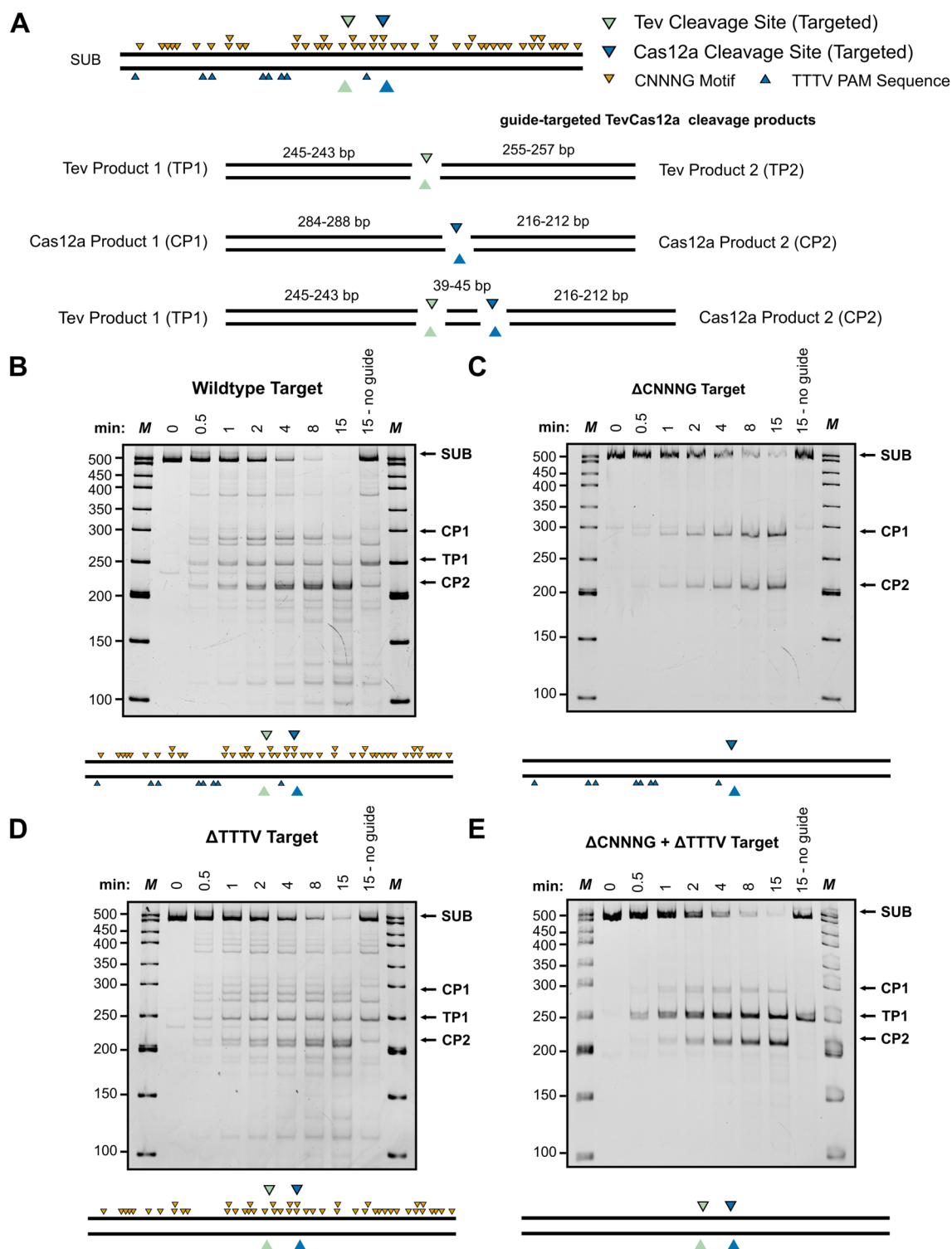


Fig. 3. TevCas12a cleavage of wildtype and mutated AAVS1 substrates. **(A)** Schematic of the linear wild-type AAVS1 substrate (SUB) indicating gRNA targeted Tev and Cas12a cleavage sites and non-targeted Tev CNNNG cleavage motifs and Cas12a TTTV PAM sequences. For simplicity, CNNNG motifs are depicted on one strand and TTTV PAM sites on the other strand. Predicted guide-targeted cleavage products are shown below. **(B–E)** TevCas12a cleavage timecourse assays on wild-type, Δ CNNNG, Δ TTTV and Δ CNNNG/ Δ TTTV substrates. For each gel image, unreacted substrate (SUB) and targeted Tev (TP1) and Cas12a (CP1-2) cleavage products are annotated. Uncropped gel images for panels B–E are shown in Supplementary Fig. S9.

Cas12a, SaCas9 displayed differential cleavage activity at the mismatched target sites relative to the on-target site at salt concentrations below 128 mM. At concentrations above 128 mM, increased fidelity towards on-target cleavage was observed.

Non-targeted cleavage by TevSaCas9, Tev-ZFE and MegaTevs

We were also interested in determining if the NTCPs were a unique feature of the TevCas12a architecture, or whether cleavage activities of other Tev-based editors could be modulated by changing *in vitro* salt concentrations. To this end, we used previously constructed Tev-gene editors; MegaTevs, a fusion of Tev to the I-OnuI meganuclease¹⁸; Tev-ZFE, a fusion of Tev to the RyA zinc finger¹⁶; and TevSaCas9, a Tev fusion to the Cas9 nuclease from *Staphylococcus aureus*²⁰ (Fig. 7A–C). To target TevSaCas9, we synthesized an SaCas9 gRNA that targeted the λ N gene. When λ DNA was incubated with the Tev-ZFE, MegaTev or TevSaCas9/gRNA $_{\lambda}$ in buffer containing 50 mM sodium chloride and 1 $\times$ rCutSmart buffer, NTCPs were observed for all three fusions (Fig. 7D–F). A product consistent with gRNA-directed cleavage by the TevSaCas9 fusion was visible (cleavage product, CP) (Fig. 7F). No discrete, targeted cleavage products were visible in reactions with the MegaTev or Tev-ZFE fusions (Fig. 7D and 7E), consistent with the fact that the meganuclease and RyA zinc finger component of the fusions were not engineered to cleave the λ genome. Mapping the cleavage products for the three editors using Oxford Nanopore revealed a consensus Tev cleavage motif for all three fusions (Fig. 7D–F).

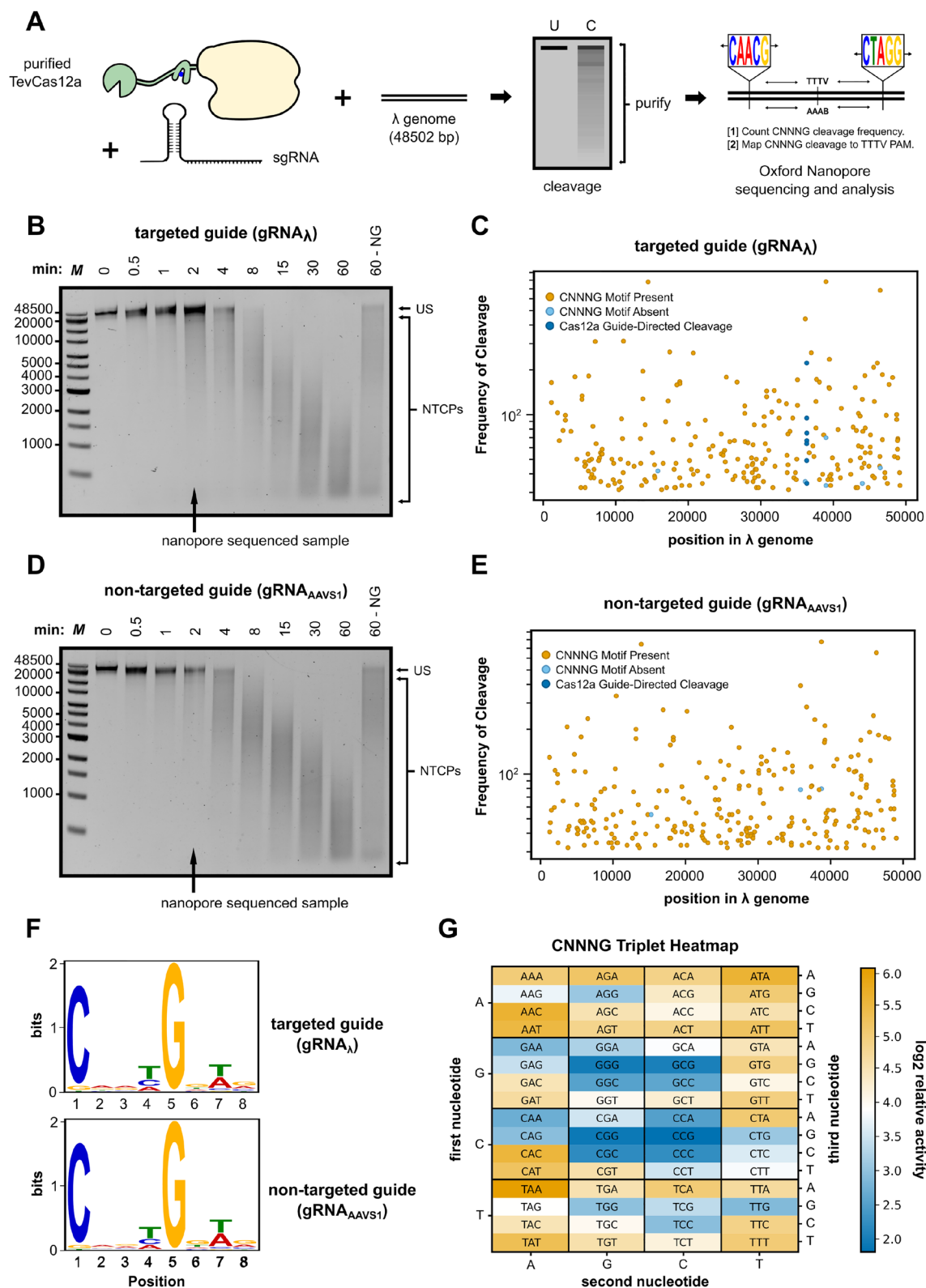
Collectively, this data shows that the activity and specificity of gene editors that use RNA- and protein-based targeting mechanisms can be modulated *in vitro* by simply changing the buffer salt concentrations. Lower salt conditions promote stability of the chimeric editors at sub-optimal target sites, enabling Tev cleavage at appropriately positioned CNNNG motifs (Fig. 7G). The cleavage preference of the Tev nuclease domain does not change regardless of the fusion context. This indicates that the effect of lowering the salt concentration is to change the stability of the Tev domain on DNA, rather than to change or relax sequence preference.

Discussion

Our *in vitro* characterization of TevCas12a activity revealed an unanticipated result; changing buffer salt concentrations effectively decoupled Tev cleavage from Cas12a-gRNA targeting and resulted in the appearance of non-gRNA directed cleavage that mapped to Tev CNNNG motifs. This decoupling was also observed for other Tev-based editors that use RNA- and protein-based targeting mechanisms. Beyond the immediate biochemical findings, de-coupling of cleavage activity from DNA targeting provides a straightforward workflow to profile the cleavage preference and specificity of nuclease domains with low DNA-binding affinity. This is particularly relevant for the GIY-YIG superfamily that is vastly underexplored in terms of cleavage preferences: of the many predicted GIY-YIG nuclease domains, only I-TevI and I-BmoI have been rigorously characterized^{31,36–38}. The ability to tether stand-alone nuclease domains to a DNA-binding module like Cas12a enables interrogation of cleavage preferences in a systematic and high-throughput fashion. Because the Cas12a scaffold can localize the fusion protein to DNA substrates without requiring sequence-specific targeting, it serves as a biochemical “anchor” to position weak-binding nuclease domains at a variety of diverse targets, providing a unique opportunity to expand the functional sequence space of GIY-YIG domains, or other nucleases.

GIY-YIG nucleases are compact modular domains that are found in a variety of protein scaffolds^{30,39–41}. The Tev nuclease domain (minimally residues 1–92) has very low affinity for DNA and an exact determination of binding affinity has been hindered by the fact that the wild-type nuclease domain (or full-length I-TevI) is extremely toxic in *E. coli* and recalcitrant to cloning and overexpression^{29,42}. Studies with the nuclease domain from the isoschizomer I-BmoI revealed a binding affinity of 10 μ M; a longer construct, encompassing residues 1–154 and analogous to the Tev 1–169 construct used here, had an affinity of 500 nM³⁷. While these measurements—and the 1.1 nM dsDNA binding constant reported for Cas12a²⁷—were obtained at different salt concentrations, they nonetheless highlight the much weaker affinity of the Tev nuclease domain. Thus, our data indicate that non-targeted cleavage products observed with TevCas12a (and other editors) arise through DNA binding domain-mediated DNA interactions. While PAM recognition enhances cleavage at proximal CNNNG motifs, transient or non-canonical DNA contacts by Cas12a are still sufficient to stabilize Tev binding and enable cleavage at more distal CNNNG sites. Indeed, previous studies with Tev-zinc finger, Tev-meganuclease, and Tev-SpCas9 fusions showed that Tev cleavage was dependent on targeted DNA-binding^{16,18,19}. Moreover, when expressed as a component of dual-nuclease editors such as MegaTev or TevCas9, the Tev nuclease domain has not displayed toxicity in either *E. coli* or human cells, as would be expected if Tev were non-specifically interacting with and cleaving DNA¹⁸.

The sensitivity of protein-DNA interactions to *in vitro* buffer conditions is well known.^{7–9} For instance, the rate of non-specific cleavage by FokI-based zinc finger nucleases can be enhanced in low salt buffers¹⁰, and the interactions of SpCas9 with non-specific sites is not impacted by changes in salt concentration^{13,14}. For TevCas12a, low sodium chloride concentrations resulted in the appearance of cleavage products in 1 $\times$ CutSmart-buffered reactions not directed by Cas12a/gRNA binding. This de-coupling of DNA cleavage from gRNA targeting is consistent with studies showing that Cas12a scans for TTTV PAM sites by facilitated diffusion, and that the rate of diffusion slows at lower salt concentrations^{15,33}. We found that non-targeted cleavage products are enhanced in low salt buffer, consistent with reduced K_{off} values that stabilize Tev–DNA complexes so that K_{cat} and K_{off} are of similar magnitude, enabling productive cleavage. In this context, Cas12a samples DNA for PAM sites through facilitated diffusion, in which sliding and hopping events along DNA increase the efficiency of target search relative to three-dimensional collisions driven by Brownian motion alone, extending local residence time for Tev cleavage at CNNNG motifs. Under low-salt conditions, reduced ionic competition at the DNA interface (‘ionic thermodynamic buffering’) stabilizes weak protein–DNA interactions, extending the dwell time of Cas12a–Tev complexes and promoting Tev cleavage at CNNNG motifs not strictly coupled to Cas12a gRNA targeting. The positional dependence of CNNNG cleavage sites within a defined window upstream of TTTV



PAMs also supports the idea that Tev cleavage is linked to diffusion of Cas12a along substrate. Tev non-targeted cleavage products are reduced at > 100 mM sodium chloride, presumably due to a faster facilitated Cas12a diffusion rate, decreased ionic thermodynamic buffering at sub-optimal Cas12a sites, and the inherently weak DNA affinity of the Tev nuclease domain. Given that all reactions contained an rCutSmart contribution of 50 mM K⁺, these conditions correspond to total monovalent cation concentrations above 150 mM. Low (12 mM) or high (102 mM) NaCl buffer did not change Tev cleavage preference; rather, it only modulates the cleavage kinetics. The fact that CNNNG motif preference remains conserved across different fusion partners—Cas12a, Cas9, meganucleases, and zinc fingers—underscores the biochemical autonomy of the Tev domain and mirrors the modularity of GIY-YIG homing endonucleases observed in nature. This modular behavior reinforces the

Fig. 4. Mapping TevCas12a non-targeted cleavage using Oxford Nanopore sequencing. **(A)** Schematic overview of the experimental workflow. Purified TevCas12a and guide RNA (sgRNA) were incubated with linear λ DNA, followed by cleavage and purification of NTCPs. Sequenced DNA fragments were mapped by position, surrounding sequence was extracted, and cleavage events were counted and measured relative to TTTV PAMs. **(B,D)** Agarose gel of TevCas12a λ genome cleavage time course with a targeted (gRNA $_{\lambda}$) or non-targeted (gRNA $_{AAVS1}$) guide. US: uncleaved substrate; NTCPs: non-target cleavage products. **(C,E)** Scatterplot depicting the frequency and distribution of top 250 cleavage events across the λ genome. Each dot represents a cleavage position, colored by CNNNG motif presence or predicted Cas12a-directed cleavage. **(F)** Sequence logos representing motifs surrounding non-targeted cleavage sites generated with the targeting and non-targeting guide RNAs. **(G)** CNNNG triplet heatmap detailing frequency of cleavage at each central triplet sequence across λ genome, normalized to genomic triplet abundance. Uncropped gel images for panels B–D are shown in Supplementary Fig. S10.

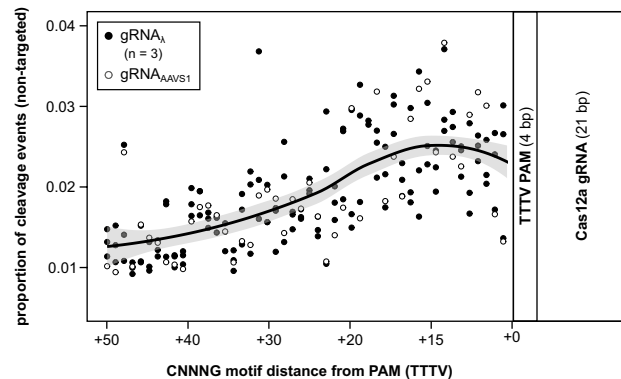


Fig. 5. Cleavage of Tev CNNNG motifs decreases with distance from the TTTV PAM. Plot of the frequency of cleavage for mapped CNNNG motifs upstream (+ bp) of the TTTV PAM. Points are from four independent mapping experiments with targeted (gRNA $_{\lambda}$) or non-targeted (gRNA $_{AAVS1}$) guide RNAs. Shaded area indicates the 95% confidence interval for the fit of the data (non-parametric smoothing function).

concept that GIY-YIG domains can be re-deployed in new architectures while retaining intrinsic biochemical features, supporting their further exploration as versatile genome editing components.

Although the primary motivation of this study was to develop *in vitro* methodologies to map the preference for Tev cleavage, a question arising from our study is the *in vivo* relevance of non-targeted cleavage by Tev-based chimeric nucleases. In this regard, TevSpCas9 cleaves at targeted sites in mammalian cell lines^{19,43} and in diatoms^{44,45}. Cas12a and SaCas9 also cleave at targeted sites in cells and are not cytotoxic. The tolerance of Cas12a and SaCas9 to target DNA mismatches in low salt conditions is consistent with stabilization of protein-DNA (or RNA-DNA) interactions that, in the context of a chimeric editor, facilitate Tev interrogation for CNNNG motifs. The opposing idea, that the Tev nuclease domain cleaves non-specifically and independently of DNA-binding domain tethering at CNNNG motifs, is not supported by the biochemical properties of the Tev nuclease domain or the observed activities of Tev-based chimeric editors where targeting is a function of the DNA-binding domain. Our data point to a remarkably consistent and specific preference of the Tev nuclease for CNNNG motifs across different fusion contexts and buffer conditions.

In summary, we have shown that changes to salt concentrations *in vitro* can modulate the cleavage activity of chimeric gene editors to de-couple nuclease domain cleavage from DNA targeting. This effect was seen for editors that use RNA- or DNA-based targeting mechanisms. From a practical standpoint, our data suggest that *in vitro* buffer conditions should be carefully considered when profiling the specificity of editors based on CRISPR nucleases, as slight changes in salt concentration (or other buffer components) could impact the activity and the readout of the tolerance to mismatches in the target site. In theory, this approach could profile the preference of any nuclease fused to a targeting domain^{46,47}, or to determine the sequence preferences of other DNA modifying domains, including base editors⁴⁸ or methylases⁴⁹. This approach would also enable mutational dissection of the Tev nuclease domain and linker to explore how specific residues contribute to cleavage-site selection, motif recognition, or catalytic activity, potentially informing the design of synthetic nucleases with tailored specificity.

Materials and methods

Bacterial strains

Escherichia coli NEB5 α F $^+$ (F $^+$ *proA*+B+ *lacI*_q δ (*lacZ*)M15zzf :: *Tn10*(*TetR*) / *fhuA2* δ (*argF* – *lacZ*)U169*phoA*glnV44 ϕ 80 δ (*lacZ*)M15gyrA96*recA1*relA1endA1*thi* – 1*hsdR17*) and *E. coli* ER2566 F $^-$ *mcrA* (*mrr-hsdRMS-mcrBC*) 80*lacZ* M15 *lacX74* *recA1* *ara139* (*ara-leu*)7697 *galU* *galK* *rpsL* (*StrR*) *endA1* *nupG* (NEB) were used for cloning and protein expression and were grown in LB media.

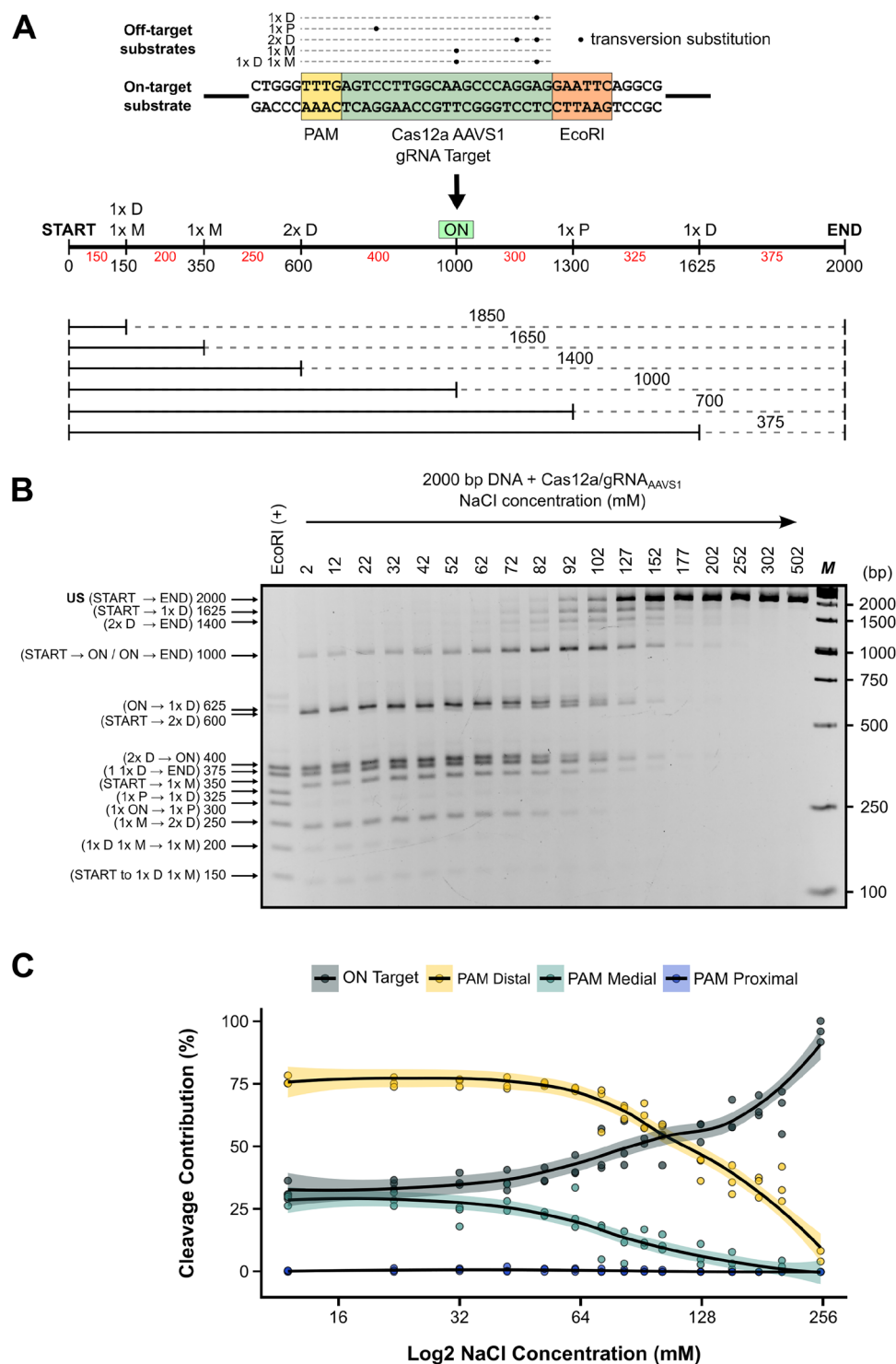


Fig. 6. Cas12a cleavage at on-target and mismatched target sites is sensitive to salt concentration. **(A)** Schematic of the linear substrate with the on-target site flanked by an EcoRI site indicated. The position of mismatches to the on-target site are indicated above the sequence, and labeled by their position in the linear substrate shown below. The anticipated cleavage products and sizes are also indicated on the linear substrate. **(B)** Representative agarose gel of Cas12a cleavage under the indicated NaCl concentrations. Sizes and identity of products are indicated on the left of the gel. US, uncut substrate. **(C)** Plot of cleavage of on-target or mismatched target sites grouped by the position of the mismatches. Points are three independent replicates with the shaded area indicated the 95% confidence interval for fit of the data to a non-parametric smoothing function. Uncropped gel image for panel B is shown in Supplementary Fig. S11.

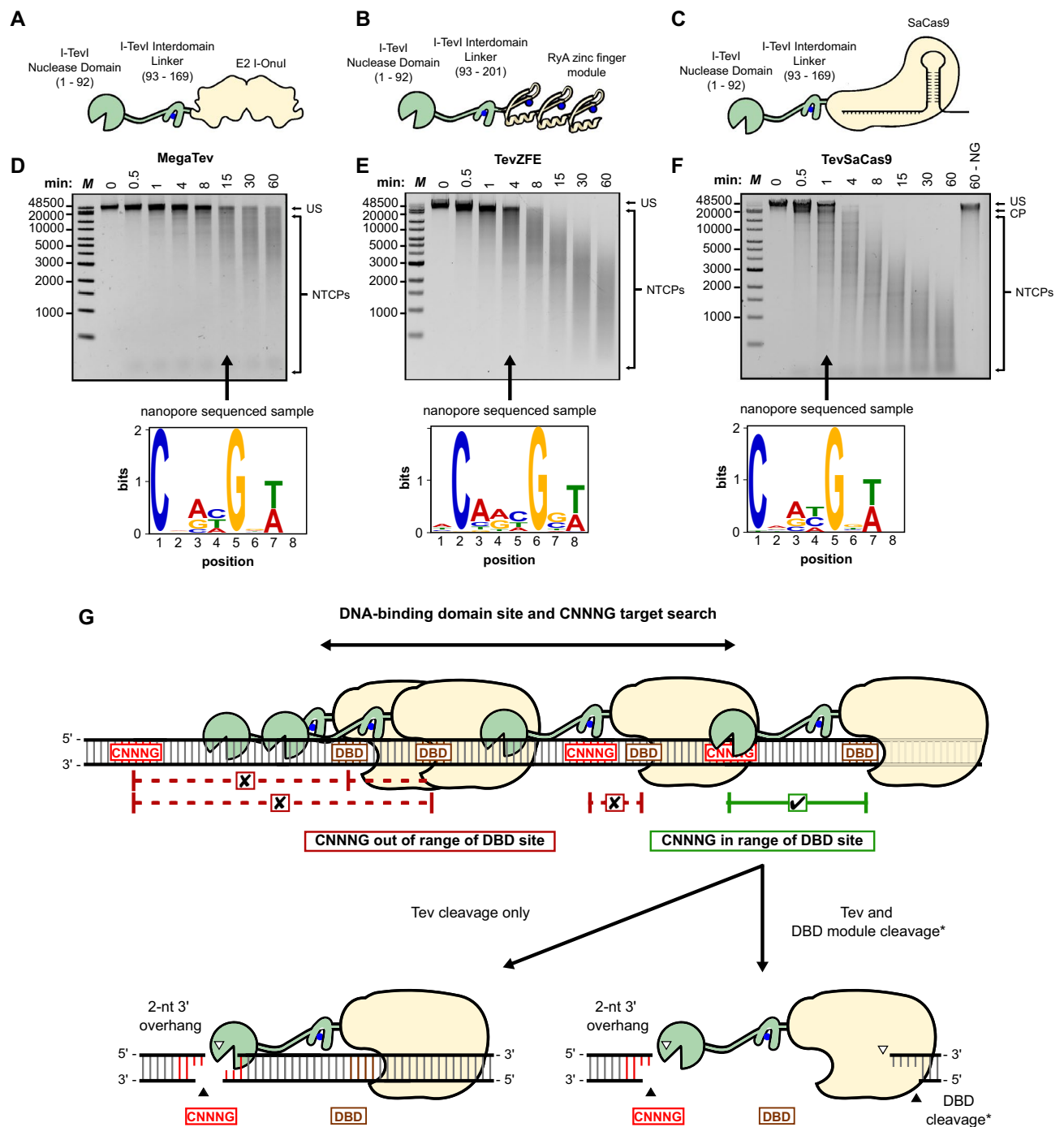


Fig. 7. Non-targeted cleavage products are observed with other chimeric gene editors. (A–C) Shown are schematics of Tev nuclease domain fusions to different DNA-binding domains and (D–F) representative time-course cleavage assays of each I-TevI fusion with λ DNA. The time point used for Oxford Nanopore sequencing is indicated. The sequence logos identified for each I-TevI fusion are shown below each gel image. US; unreacted substrate; NTCs; non-targeted cleavage products. (G) Model of non-targeted cleavage by chimeric Tev-based nucleases. Tev editors engage with DNA substrate and scan for target sites (DBD). If engaged at a DBD and a Tev CNNNG motif is in range (green), cleavage occurs at the Tev and DBD module sites (in the case of active TevCas fusion proteins). Cleavage by Tev does not proceed if Tev editors are engaged at DBD sites where the CNNNG motif is out of range (red). Uncropped gel images for panels D–F are shown in Supplementary Fig. S12.

Plasmid construction

All TevCas12a variants were cloned into the pET11a expression vector using Gibson Assembly with NEBuilder[®] HiFi, following the manufacturer's protocol. To create TevCas12a fusions, *E. coli* codon-optimized I-TevI (amino acids 1–169, GenBank: P13299.2) with a GGSGGTGGSG peptide at the C-terminal end was fused to the N-terminus of AsCpf1 (Cas12a, GenBank: U2UMQ6.1). Mutations that inactivated I-TevI (R27A) or Cas12a (D908P) were introduced into the wild-type TevCas12a plasmid using golden gate mutagenesis. Primers for mutagenesis and cloning are listed in Table S1. All constructs were verified by Sanger sequencing of the cloned insert (London Regional Genomics Facility, London, ON) or whole plasmid sequencing (Plasmidsaurus). A list of oligonucleotides used in cloning and plasmids used in this study are provided in Table S1 and S2, respectively.

Guide RNA synthesis

All guide RNA sequences and oligonucleotides are listed in Table S1. Guide RNAs were synthesized by one-pot *in vitro* transcription using the T7 RiboMAX[®] Express Large Scale RNA Production System (Promega). DNA templates were prepared using two distinct strategies depending on the nuclease. For Cas12a gRNAs, templates were formed by annealing fully complementary sense and antisense oligonucleotides (2 μ M each; Integrated DNA Technologies), yielding a 69 bp duplex encoding a T7 promoter, spacer, gRNA scaffold, and a 3' poly-T terminator. No template extension was required. For SaCas9 gRNAs, templates were generated by annealing a guide-specific forward oligo (containing the T7 promoter, spacer, and partial scaffold) with a universal reverse oligo (DE6018) encoding the remainder of the scaffold and a T-rich terminator. These partially overlapping oligos were extended *in situ* using Klenow Fragment (3'→5' exo-, New England Biolabs). Final 20 μ L reactions contained 2.5 U Klenow, 125 μ M dNTPs, 1× RiboMAX Express T7 Buffer, 2 μ L T7 Express Enzyme Mix (Promega), and the annealed DNA duplex. Reactions were incubated at 37°C for 4 h. Transcription products were purified using the Monarch[®] RNA Cleanup Kit (New England Biolabs) according to the manufacturer's protocol. No DNase treatment was performed. RNA integrity was assessed by agarose gel electrophoresis.

DNA substrates for *in vitro* assays

AAVS1 substrates were amplified from HEK293 genomic DNA using PrimeSTAR GXL DNA polymerase (Takara Bio) and primers OL0345 and OL0346. The 872 bp amplicon corresponds to the AAVS1 locus (GenBank: AC010327.8) and was cloned into the pJET vector (pJET-AAVS1) using the ThermoFisher CloneJET PCR Cloning Kit. Sequence integrity was confirmed by Sanger sequencing of the genomic PCR product, and again following cloning.

A linear Cas9/Cas12a off-target DNA substrate ("Cas-OFF") was assembled by BsaI-HFv2 Golden Gate cloning of five gBlocks (Integrated DNA Technologies), each flanked by BsaI recognition sites. Equimolar amounts (0.05 pmol) were mixed and ligated using BsaI-HFv2 (New England Biolabs), T4 DNA Ligase (New England Biolabs), and 1× T4 DNA ligase buffer (New England Biolabs). The assembled product was used directly as template for PCR amplification with primers DE7182 and DE7184.

Substrates lacking CNNNG and TTTV motifs were PCR-amplified from gBlocks (IDT) using gBlocks as templates with primers DE6502 through DE6785. Lambda DNA was purchased from New England Biolabs.

Protein purification

C-terminal 6×His-tagged TevCas12a variants and TevSaCas9 (in pET11a), as well as MegaTev and TevRyA(ZFE) (in pACYC-Duet1), were expressed under the control of a T7 promoter in *E. coli* ER2566 cells. Cultures (250 mL) were grown in a 2:1 flask-to-volume ratio at 37°C to an OD₆₀₀ of 0.6, then induced with 1 mM IPTG and incubated overnight at 16°C.

Cells were harvested by centrifugation and lysed by sonication in binding buffer [20 mM Tris HCl, pH 8.0, 500 mM NaCl, 1 mM DTT, 5% (v/v) glycerol, 10 mM imidazole] supplemented with 1 mM phenylmethylsulfonyl fluoride (PMSF). Clarified lysates were applied to 1 mL HisTrap[™] Ni-NTA affinity columns (Cytiva), and bound proteins were eluted in a stepwise gradient of imidazole (100, 200, 300, and 500 mM). Eluted fractions were dialyzed using 8 kDa MWCO tubing into storage buffer (identical to binding buffer but lacking imidazole), aliquoted, flash-frozen in a –80°C ethanol bath, and stored at –80°C.

In vitro cleavage reactions

Cleavage reactions were performed in biological triplicate at varying molar ratios of RNP to substrate (20:1, 10:1, 5:1, 1:1, or 1:5) using TevCas12a RNPs, TevSaCas9 RNPs, MegaTev, or TevRyA(ZFE). RNPs were assembled at a 1:1.5 molar ratio of protein to guide RNA by incubation at room temperature for 10 minutes. Reactions contained 2.5 nM DNA substrate in 50 mM NaCl, 20 mM Tris HCl (pH 8.0), 1 mM DTT, and 1× rCutSmart buffer (New England Biolabs; 50 mM potassium acetate, 20 mM Tris-acetate, 10 mM magnesium acetate, 100 μ g/mL BSA), and were incubated at 18°C, 24°C, 30°C, or 37°C. Aliquots were taken at designated time points and quenched in stop solution (16.7 mM EDTA, 50 μ g/mL RNase A, 50 μ g/mL proteinase K), followed by incubation at 37°C for 30 minutes.

For salt titration experiments, TevCas12a, Cas12a and SaCas9 RNPs were assembled and diluted to 500 nM in nuclease-free water to minimize salt carryover from the storage buffer. NaCl was titrated (0–500 mM final) into reaction mixtures containing 2.5 nM λ DNA or a 2000 bp linear "Cas-OFF" amplicon in 20 mM Tris HCl (pH 8.0), 1 mM DTT, and 1× rCutSmart buffer. RNPs were added last at a 20:1 RNP:DNA molar ratio, and reactions were incubated at 37°C for 1 hour (TevCas12a) or 15 minutes (Cas12a, SaCas9) before quenching.

Cleavage products were analyzed by electrophoresis on 12% (v/v) polyacrylamide gels in Tris-borate-EDTA (TBE) buffer for linear substrates, or 0.8% (w/v) agarose gels in Tris-acetate-EDTA (TAE) buffer for pJET-AAVS1 and λ DNA. Gel images were analyzed in ImageLab (BioRad) by manual band calling based on fragment size, and cleavage fractions were determined from relative band intensities.

Oxford nanopore sequencing and data analysis

Pooled cleavage reactions containing 1.5 µg of pJET-AAVS1 or λ DNA were purified using the Monarch PCR & DNA Cleanup Kit (New England Biolabs) according to the manufacturer's protocol. End-repair and dA-tailing were performed using the NEBNext[®] Companion Module, followed by ligation of barcoded adapters from the Native Barcoding Expansion kit (EXP-NB104, Oxford Nanopore Technologies). Ligations were incubated at room temperature for 30 minutes and purified using AMPure XP beads. A total of 800 ng of barcoded product was recovered and split equally across samples. Final pooled input for sequencing was 500 ng. Library preparation followed the SQK-LSK114 protocol. DNA concentration and integrity were assessed using a Qubit fluorometer and agarose gel electrophoresis.

Sequencing was performed on MinION flow cells (R9.4.1), yielding 2–3 million reads per run. Basecalling and demultiplexing were performed with Guppy (v6.5.7) using default parameters. Reads were aligned to the corresponding reference sequences (pJET-AAVS1 and λ DNA) using minimap2 with the `-ax map-ont` preset. Mapped reads were parsed using a custom Perl script to extract read length, strand orientation, and mapped position. Cleavage sites were defined as the mapped termini of reads relative to the reference sequence.

Custom Python scripts were used to identify cleavage positions and characterize sequence features at, upstream, and downstream of each site. Cleavage frequencies by position were calculated, and recurrent motifs and PAM proximities were identified. Motif logos were generated using MEME (v5.5.5) with the "one occurrence per sequence" (oops) model.

To assess proximity to protospacer-adjacent motifs, each cleavage site was scanned for the nearest upstream TTTV [0] PAM (defined by end coordinate and strand mapping), and the nearest downstream or BAAA [16] PAM (defined by start coordinate and strand mapping). Linear distances between cleavage positions and PAM sites were calculated. All analyses were performed using custom Python scripts provided in Supplementary Information.

Data availability

The Oxford Nanopore sequences generated in this study have been deposited in the Sequence Read Archive [PRJNA1273990]. Computer code is provided in the Supplementary Information.

Received: 9 July 2025; Accepted: 30 October 2025

Published online: 28 November 2025

References

- Carroll, D. Genome engineering with targetable nucleases. *Annu. Rev. Biochem.* **83**, 409–439 (2014).
- Kim, H. & Kim, J.-S. A guide to genome engineering with programmable nucleases. *Nat. Rev. Genet.* **15**, 321–334 (2014).
- Pattanayak, V., Ramirez, C. L., Joung, J. K. & Liu, D. R. Revealing off-target cleavage specificities of zinc-finger nucleases by in vitro selection. *Nat. Methods* **8**, 765–770 (2011).
- Pattanayak, V. et al. High-throughput profiling of off-target dna cleavage reveals rna-programmed cas9 nuclease specificity. *Nat. Biotechnol.* **31**, 839–843 (2013).
- Zhang, L. et al. Systematic in vitro profiling of off-target affinity, cleavage and efficiency for CRISPR enzymes. *Nucleic Acids Res.* **48**, 5037–5053 (2020).
- Winter, R. B., Berg, O. G. & Von Hippel, P. H. Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The Escherichia coli Lac repressor-operator interaction: kinetic measurements and conclusions. *Biochemistry* **20**, 6961–6977 (1981).
- Berg, O. G., Winter, R. B. & Von Hippel, P. H. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry* **20**, 6929–6948 (1981).
- Lohman, T. M. & Von Hippel, P. H. Kinetics of protein-nucleic acid interactions: Use of salt effects to probe mechanisms of interaction. *Crit. Rev. Biochem.* **19**, 191–245 (1986).
- Bonnet, I. et al. Sliding and jumping of single EcoRV restriction enzymes on non-cognate DNA. *Nucleic Acids Res.* **36**, 4118–4127 (2008).
- Schierling, B. et al. A novel zinc-finger nuclease platform with a sequence-specific cleavage module. *Nucleic Acids Res.* **40**, 2623–2638 (2012).
- Rinaldi, F. C., Doyle, L. A., Stoddard, B. L. & Bogdanove, A. J. The effect of increasing numbers of repeats on TAL effector dna binding specificity. *Nucleic Acids Res.* **45**, 6960–6970 (2017).
- Cuculis, L. et al. Divalent cations promote tale dna-binding specificity. *Nucleic Acids Res.* **48**, 1406–1422 (2020).
- Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C. & Doudna, J. A. DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature* **507**, 62–67 (2014).
- Yang, M. et al. Nonspecific interactions between SpCas9 and dsDNA sites located downstream of the PAM mediate facilitated diffusion to accelerate target search. *Chem. Sci.* **12**, 12776–12784 (2021).
- Jeon, Y. et al. Direct observation of DNA target searching and cleavage by CRISPR-Cas12a. *Nat. Commun.* **9**, 2777 (2018).
- Kleinstiver, B. P. et al. Monomeric site-specific nucleases for genome editing. *Proc. Natl. Acad. Sci.* **109**, 8061–8066 (2012).
- Kleinstiver, B. P. et al. The I-TevI nuclease and linker domains contribute to the specificity of monomeric TALENs. *G3: Genes, Genomes, Genet.* **4**, 1155–1165 (2014).
- Wolfs, J. M. et al. Megatevs: single-chain dual nucleases for efficient gene disruption. *Nucleic Acids Res.* **42**, 8816–8829 (2014).
- Wolfs, J. M. et al. Biasing genome-editing events toward precise length deletions with an RNA-guided TevCas9 dual nuclease. *Proc. Natl. Acad. Sci.* **113**, 14988–14993 (2016).
- Hamilton, T. A. et al. Efficient inter-species conjugative transfer of a CRISPR nuclease for targeted bacterial killing. *Nat. Commun.* **10**, 4544 (2019).
- Zetsche, B. et al. Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system. *Cell* **163**, 759–771 (2015).
- Swarts, D. C. Making the cut (s): How Cas12a cleaves target and non-target DNA. *Biochem. Soc. Trans.* **47**, 1499–1510 (2019).
- Klener, M. J., Koob, J. G., Gootenberg, J. S., Abudayyeh, O. O. & Zhang, F. SHERLOCK: nucleic acid detection with CRISPR nucleases. *Nat. Protoc.* **14**, 2986–3012 (2019).
- Paul, B. & Montoya, G. CRISPR-Cas12a: Functional overview and applications. *Biomed. J.* **43**, 8–17 (2020).
- Ma, E. et al. Improved genome editing by an engineered CRISPR-Cas12a. *Nucleic Acids Res.* **50**, 12689–12701 (2022).
- Li, S.-Y. et al. CRISPR-Cas12a-assisted nucleic acid detection. *Cell Discovery* **4**, 20 (2018).
- Strohkend, I., Saifuddin, F. A., Rybarski, J. R., Finkelstein, I. J. & Russell, R. Kinetic basis for DNA target specificity of CRISPR-Cas12a. *Mol. Cell* **71**, 816–824 (2018).

28. Murugan, K., Seetharam, A. S., Severin, A. J. & Sashital, D. G. CRISPR-Cas12a has widespread off-target and dsDNA-nicking effects. *J. Biol. Chem.* **295**, 5538–5553 (2020).
29. Van Roey, P., Meehan, L., Kowalski, J. C., Belfort, M. & Derbyshire, V. Catalytic domain structure and hypothesis for function of GIY-YIG intron endonuclease I-TevI. *Nat. Struct. Biol.* **9**, 806–811 (2002).
30. Kowalski, J. C. et al. Configuration of the catalytic GIY-YIG domain of intron endonuclease I-TevI: coincidence of computational and molecular findings. *Nucleic Acids Res.* **27**, 2115–2125 (1999).
31. Edgell, D. R., Stanger, M. J. & Belfort, M. Coincidence of cleavage sites of intron endonuclease I-TevI and critical sequences of the host thymidylate synthase gene. *J. Mol. Biol.* **343**, 1231–1241 (2004).
32. Dean, A. B. et al. Zinc finger as distance determinant in the flexible linker of intron endonuclease I-TevI. *Proc. Natl. Acad. Sci.* **99**, 8554–8561 (2002).
33. Singh, D. et al. Real-time observation of DNA target interrogation and product release by the RNA-guided endonuclease CRISPR Cpf1 (Cas12a). *Proc. Natl. Acad. Sci.* **115**, 5444–5449 (2018).
34. Kim, D. et al. Genome-wide analysis reveals specificities of Cpf1 endonucleases in human cells. *Nat. Biotechnol.* **34**, 863–868 (2016).
35. Kleinstiver, B. P. et al. Genome-wide specificities of CRISPR-Cas Cpf1 nucleases in human cells. *Nat. Biotechnol.* **34**, 869–874 (2016).
36. Bryk, M. et al. The td intron endonuclease i-tevi makes extensive sequence-tolerant contacts across the minor groove of its dna target. *EMBO J.* **12**, 2141–2149 (1993).
37. Kleinstiver, B. P., Wolfs, J. M. & Edgell, D. R. The monomeric GIY-YIG homing endonuclease I-BmoI uses a molecular anchor and a flexible tether to sequentially nick DNA. *Nucleic Acids Res.* **41**, 5413–5427 (2013).
38. Edgell, D. R., Stanger, M. J. & Belfort, M. Importance of a single base pair for discrimination between intron-containing and intronless alleles by endonuclease i-bmoi. *Curr. Biol.* **13**, 973–978 (2003).
39. Dunin-Horkawicz, S., Feder, M. & Bujnicki, J. M. Phylogenomic analysis of the GIY-YIG nuclease superfamily. *BMC Genomics* **7**, 1–19 (2006).
40. Sharma, M., Ellis, R. L. & Hinton, D. M. Identification of a family of bacteriophage T4 genes encoding proteins similar to those present in group I introns of fungi and phage. *Proc. Natl. Acad. Sci.* **89**, 6658–6662 (1992).
41. Chenal, M. et al. Discovery of the widespread site-specific single-stranded nuclease family Ssn. *Nat. Commun.* **16**, 2388 (2025).
42. Derbyshire, V., Kowalski, J. C., Dansereau, J. T., Hauer, C. R. & Belfort, M. Two-domain structure of the td intron-encoded endonuclease I-TevI correlates with the two-domain configuration of the homing site. *J. Mol. Biol.* **265**, 494–506 (1997).
43. Kelly, R. D. et al. Noncanonical functions of Ku may underlie essentiality in human cells. *Sci. Rep.* **13**, 12162 (2023).
44. Slattery, S. S. et al. An expanded plasmid-based genetic toolbox enables Cas9 genome editing and stable maintenance of synthetic pathways in *Phaeodactylum tricornutum*. *ACS Synth. Biol.* **7**, 328–338 (2018).
45. Slattery, S. S. et al. Plasmid-based complementation of large deletions in *Phaeodactylum tricornutum* biosynthetic genes generated by Cas9 editing. *Sci. Rep.* **10**, 13879 (2020).
46. Guilinger, J. P., Thompson, D. B. & Liu, D. R. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* **32**, 577–582 (2014).
47. Tsai, S. Q. et al. Dimeric CRISPR RNA-guided FokI nucleases for highly specific genome editing. *Nat. Biotechnol.* **32**, 569–576 (2014).
48. Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A. & Liu, D. R. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* **533**, 420–424 (2016).
49. Vojta, A. et al. Repurposing the CRISPR-Cas9 system for targeted DNA methylation. *Nucleic Acids Res.* **44**, 5615–5628 (2016).

Acknowledgements

The authors thank members of the Edgell lab and Specific Biologics Inc. for comments on the manuscript.

Author contributions

K.W.L., A.L.W., T.A.M., B.E.S. and D.R.E. conceived the experiments. K.W.L., T.A.M., B.E.S. and A.L.W. conducted the experiments. K.W.L., A.L.W. and D.R.E. analyzed the results. K.W.L. and D.R.E. wrote the manuscript. K.W.L., A.L.W., T.A.M., B.E.S. and D.R.E. read the manuscript.

Funding

This work was supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada to D.R.E. [RGPIN-2022-05459] an Alliance Grant from the Natural Sciences and Engineering Research Council of Canada and MITACS Canada to D.R.E. and B.E.S. [ALLRP 571374 - 21]. K.W.L. was supported by an Ontario Graduate Scholarship from the Government of Ontario.

Declarations

Competing interests

B.E.S. and T.A.M. are employees of Specific Biologics Inc. D.R.E. is co-founder of Specific Biologics Inc. All the remaining authors (K.W.L. and A.L.W.) declare no conflict of interest.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-26759-7>.

Correspondence and requests for materials should be addressed to D.R.E.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025