

TITLE

Rational design of an enhanced genome editing protein for therapeutic gene correction and diagnostics

INTRODUCTION

RNA-guided, programmable CRISPR-Cas9 nucleases which is native to prokaryotes as a part of the adaptive immune system, is a facile repurposed genome editing system heralding a new dawn in the field of therapeutic gene correction and disease detection. Like most other genome editing tools, CRISPR-based editing too depends on the generation of a double-strand break (DSB) at a defined position in the genome and its subsequent cellular repair to mend the break by either of the two major cellular repair pathways viz. non-homologous-end joining (NHEJ) and homology-directed repair (HDR) (Acharya, Maiti, and Chakraborty 2020). NHEJ is the most commonly used error-prone cellular repair machinery which leads to small insertions and deletions (indels) at the breakpoint. Alternatively, HDR faithfully restores the damaged sequence with a homologous sequence found either in another chromosome or externally supplied. These repair pathways have been harnessed to inactivate genes (knock-outs) or to introduce new sequences (knock-ins) and these outcomes have been used ingeniously to correct disease-causing mutation/s (Acharya, Maiti, and Chakraborty 2020). At the heart of this system lies the protein component Cas9, an RNA-guided endonuclease derived from the type II CRISPR-Cas system and its associated RNA component chimeric single-guide (sg) RNA derived from the CRISPR RNA (crRNA) and the noncoding trans-activating RNA (tracrRNA). This binary complex (ribonucleoprotein, RNP) guided by sgRNA searches its target DNA upstream of a protospacer adjacent motif (PAM) (a small DNA sequence which varies from one Cas9 protein to another) and makes a ternary complex (RNP and target DNA) leading to the final editing outcome by either of the earlier mentioned repair pathways (Acharya, Maiti, and Chakraborty 2020).

Safe therapeutic gene correction can only be accomplished if the Cas9 has high on-target editing efficiency and minimal off-target activity. The off-targeting effect of the most widely used system of Cas9 from *Streptococcus pyogenes* (SpCas9) and the low editing efficacy of its engineered derivatives raise serious concern on its utility as a therapeutic modality (Tsai et al. 2015). Recently, we reported a Cas9 from *Francisella novicida* which has negligible affinity towards off-target DNA substrate thereby giving specificity in DNA interrogation (Acharya et al. 2019). However, its low cellular targeting efficiency in the human genome is limiting its therapeutic application. We envisioned that enhancing the functional activity of FnCas9 protein might circumvent this issue. Thus, we set out to rationally engineer this FnCas9 protein guided by structure to enhance its kinetic activity by retaining its intrinsic specificity.

OBJECTIVES

- 1. Discovery of a specific genome editing protein for human genome engineering**
- 2. Development of an enhanced genome editing protein by structure-guided rational engineering**

MATERIAL AND METHODS

Plasmid construction

The gene encoding full-length *F. novicida* (Fn) Cas9 nuclease residues 1 to 1,629 bp was PCR-amplified using PX408 (Addgene 68705) as a template and cloned in pET28-His-10-Smt3 vector (a kind gift from Stewart Shuman, Molecular Biology Program Sloan-Kettering Institute, New York, and K. M. Sinha, Amity Institute of Biotechnology, Amity University, Gurgaon, Haryana, India) and pET-His6-GFP-TEV-LIC vector (Addgene 29663), following restriction enzyme-based cloning and ligation-independent cloning (LIC), respectively. Catalytically inactive FnCas9 double mutants, Point mutations on the pET-His6-dFnCas9GFP backbone and PX458-3xHA-FnCas9 backbone (Addgene 130969) were generated on

pET-His6-FnGFP-TEV-LIC plasmid backbone by a QuikChange II site-directed mutagenesis kit (Agilent) following the manufacturer's protocol, with some modifications. Point mutations and deletions were done by inverse PCR method on FnCas9 cloned in pE-Sumo vector backbone (LifeSensors) where intended changes were made on the forward primer and the entire plasmid was amplified by inverse PCR.

Protein and sgRNA purification

The proteins used in this study were purified as reported previously. Briefly, plasmids for Cas9 from *Francisella novicida* were expressed in Escherichia coli Rosetta2 (DE3) (Novagen). The protein-expressing Rosetta2 (DE3) cells were cultured at 37°C in LB medium (supplemented with 50mg/ml kanamycin) until OD₆₀₀ reached 0.6 and protein expression was induced by the addition of 0.5mM isopropyl-β-D-thiogalactopyranoside (IPTG). The Rosetta2 (DE3) cells were further cultured at 18°C overnight and harvested by centrifugation. The *E.coli* cells were resuspended in buffer A (20 mM Tris-HCl, pH 8.0, 20 mM imidazole, and 1 M NaCl), and lysed by sonication, and centrifuged. The lysate was mixed with Ni-NTA beads (Roche), the mixture was loaded into a Poly-Prep Column (BioRad) and the protein was eluted by buffer B (20 mM Tris-HCl, pH 8.0, 0.3 M imidazole, and 0.3 M NaCl). The affinity eluted protein was mixed with ion-exchange beads (SP Sepharose Fast Flow, GE Healthcare) equilibrated with buffer C (20 mM Tris-HCl, pH 8.0, and 0.15 M NaCl) and the protein was eluted by buffer D (20 mM Tris-HCl, pH 8.0, and 1 M NaCl). The concentration of purified protein was measured by the Pierce BCA protein assay kit (Thermo Fisher Scientific). The purified proteins were stored at -80 °C until further use.

In vitro transcribed sgRNAs were synthesized using MegaScript T7 Transcription kit (Thermo Fisher Scientific) using T7 promoter containing template as substrates. IVT reactions were incubated overnight at 37°C followed by NucAway spin column (Thermo Fisher Scientific) purification as described earlier (Acharya et al. 2019). IVT sgRNAs were stored at -20 °C until further use.

***in vitro* cleavage (IVC) assay**

For the kinetic study, the pUC119 plasmid containing the target sequence and the respective PAM (mentioned in respective legends) was used as the substrate for *in vitro* cleavage experiments. The linearized pUC119 plasmid (50 ng or ~5nM) was incubated at 37°C for 0.5–5 min with the Cas9–sgRNA complex (50 nM) in 10 µL of reaction buffer, containing 20 mM HEPES, pH 7.5, 150 mM KCl, 10 mM MgCl₂, 1 mM DTT, and 5% glycerol. The reaction was stopped by the addition of a quenching buffer, containing EDTA (20 mM final) and Proteinase K (40 ng). The reaction products were resolved, visualized, and quantified with a MultiNA microchip electrophoresis device (SHIMADZU) (Nishimasu et al. 2018).

The rest of the IVC assays were done as described earlier (Acharya et al. 2019). Details of substrates, concentrations, and incubation time are mentioned in respective figure legends.

PAM discovery assay

The PAM discovery assays were performed, as previously described (Nishimasu et al. 2018). Briefly, a library of pUC119 plasmids containing eight randomized nucleotides downstream of the target sequence was incubated at 37°C for 5 min with the FnCas9–sgRNA complex (50 nM), in 50 µL of the reaction buffer. The reactions were quenched by the addition of Proteinase K and then purified using a Wizard DNA Clean-Up System (Promega). The purified DNA samples were amplified for 25 cycles, using primers containing common adapter sequences. After column purification, each PCR product (~5 ng) was subjected to the second round of PCR for 15 cycles, to add custom Illumina TruSeq adapters and sample indices. The sequencing libraries were quantified by qPCR (KAPA Biosystems) and then subjected to paired-end sequencing on a MiSeq sequencer (Illumina) with 20% PhiX spike-in (Illumina). The sequencing reads were demultiplexed by primer sequences and sample indices, using NCBI Blast+ (version 2.8.1) with the blastn-short option. For each sequencing sample, the number of reads for every possible 8-nt PAM sequence pattern ($4^8 = 65,536$ patterns in total) was counted and normalized by the total number of reads in each sample. For a given PAM

sequence, the enrichment score was calculated as log₂ -fold enrichment as compared to the untreated sample. PAM sequences with enrichment scores of – 2.0 or less were used to generate the sequence logo representation using WebLogo3 (<http://weblogo.threeplusone.com/create.cgi>).

Binding assay

MST was performed as described previously (Acharya et al. 2019). Briefly, dFnCas9-GFP and variant proteins were complexed with PAGE purified respective IVT sgRNAs (purified by 12% Urea-PAGE). The binding affinities of the Cas9 proteins and sgRNA RNP complexes were calculated using Monolith NT. 115 (NanoTemper Technologies GmbH, Munich, Germany). RNP complex (Protein:sgRNA molar ratio,1:1) was reconstituted at 25 for 10 mins in reaction buffer (20 mM HEPES, pH7.5, 150mM KCl, 1mM DTT, 10mM MgCl₂) HPLC purified 30 bp dsDNA (IDT) of different genomic loci with varying concentrations (ranging from 0.09 nM to 30 µM) were incubated with RNP complex at 37° C temperature for 30 min in reaction buffer. The sample was loaded into NanoTemper standard treated capillaries and measurements were performed at 25°C using 20% LED power and 40% MST power. Data analyses were done using NanoTemper analysis software and the data were plotted by OriginLab.

Cell culture

HEK293T cells were grown in DMEM media supplemented with high glucose (Invitrogen), 2 mM GlutaMax, 10% FBS (Invitrogen), 1X antibiotic and antimycotic (Invitrogen) at 37°C in 5% CO₂. Transfections of mammalian cells were performed using Lipofectamine 3000 Reagent (Invitrogen) following the manufacturer's protocol.

Amplicon sequencing

HEK293T cells on six well dishes were transfected with 2 µg of respective Cas9 containing sgRNAs. 48 h post-transfection GFP-positive cells were sorted by BD FACS Melody (BD Biosciences). 48 hrs post-transfection GFP-positive cells were FACS sorted

(BD FACSMelody Cell Sorter) and gDNA was isolated (Lucigen QuickExtract Extraction solution). PCR primers were designed flanking the predicted double stranded break site and amplified with Phusion High-Fidelity DNA polymerase (Thermo Fisher Scientific). The 16S Metagenomic sequencing library preparation protocol was adapted for library preparation. Briefly, the respective loci was amplified using forward and reverse primers along with overhang adapter sequences using Phusion High-Fidelity DNA polymerase (Thermo Fisher). AMPure XP beads (A63881, Beckman Coulter) were used to separate out amplicons from free primers and primer dimers. Dual indexing was done using Nextera XT V2 index kit followed by a second round of bead-based purification. The libraries were quantified using a Qubit dsDNA HS Assay kit (Invitrogen, Q32853) and were also loaded on agarose gel for the qualitative check. Libraries were normalized, pooled and were loaded onto the Illumina MiniSeq platform.

HDR assay at DCX locus in HEK293T:

HEK293T cells were cultured in DMEM with GlutaMAX supplement (ThermoFisher Scientific Cat. No. 10566016) with 10% FBS serum. 70%-80% confluent HEK293T cells were harvested from a 6 well plate using Trypsin-EDTA (0.05%) (ThermoFisher Scientific Cat. No.: 25300062) and pipetted to make a single-cell suspension. For each electroporation reaction, a total 15ug plasmid was mixed in Resuspension buffer R, in which linearized donor plasmid DNA and Cas9-gRNA vector were taken in a 1:2 ratio. 6×10^5 cells were resuspended in 100 μ L of Resuspension Buffer R containing plasmids and electroporation was performed using Neon Transfection System 100 μ L Kit (ThermoFisher Scientific Cat. No. MPK10096) with double pulses at 950 V, 30 milliseconds pulse width. The electroporated cells were transferred immediately to a 6 well plate containing 2 ml of pre-warmed culture medium and incubated at 37°C and 5% CO₂. After 24 hours cells were washed and re-incubated with a fresh culture medium. 72 hours post electroporation GFP positive cells per sample were sorted using BD FACSMelody Cell Sorter (BD Biosciences-US) and gDNA was isolated from the sorted cells using Wizard Genomic DNA Purification Kit (Promega) for qPCR genotyping.

qPCR reactions were performed using LightCycler 480 SYBR Green I Master (Roche) added to 50ng DNA for each sample. The cycling conditions on the instrument were as follows: Initial denaturation 95°C for 5 min followed by 40 amplification cycles of 95°C for 10 sec; 63°C for 30 sec; 72°C for 30.

RESULTS

1. Discovery of a specific genome editing protein for human genome engineering

Towards the first objective, we set out to find a Cas9 protein which is intrinsically more specific than the most widely used Cas9 protein from *Streptococcus pyogenes* (SpCas9) which is having simpler NGG PAM recognition (Nishimasu et al. 2014). Although several Cas9 proteins recognizing different PAM sequences have so far been reported in the literature, only a subset of these have been characterized and have demonstrated genome-editing ability in eukaryotic cells (Cong et al. 2013; Esvelt et al. 2013; Mali et al. 2013; Ran et al. 2015; E. Kim et al. 2017). Cas9 from *Francisella novicida* (FnCas9) is one of the largest Cas9 orthologs which is having predominant NGG PAM recognition as SpCas9 (H. Hirano et al. 2016). The crystal structure of FnCas9 ribonucleoprotein (RNP) in complex with target DNA has revealed both conserved and divergent features of interaction that is unique among the Cas9 enzymes studied. Unlike SpCas9, FnCas9 does not form a bilobed structure, has a different sgRNA scaffold, and has been implicated in RNA targeting. Although the protein can efficiently cleave DNA substrates *in vitro*, its *in vivo* activity at several genomic loci is significantly diminished as compared to SpCas9 (H. Hirano et al. 2016). We hypothesized that the unique structural attributes might give FnCas9 a unique DNA interrogation property.

The specificity of a genome-editing protein is guided by a balance between its affinity for target and the ability to discriminate off-targets. A recent report has suggested that FnCas9 shows higher intrinsic specificity than SpCas9 to its target by showing less

tolerance to single mismatches at certain sgRNA positions (F. Chen et al. 2017). We investigated the in vitro cleavage efficiency of both SpCas9 and FnCas9 by systematically changing every base in a given substrate and observed that, whereas SpCas9 cleaved nearly equally at all mismatched positions, FnCas9 was less tolerant to single mismatches particularly at base positions 19, 18, 17, and 16 at the PAM distal end, where considerable cleavage abrogation was observed ($\sim 56 \pm 7\%$) (Fig1A). Interestingly, engineered highly specific Cas9 variants that are able to prevent cleavage at off targets containing PAM distal mismatches adopt a conformational structure that renders the HNH DNA cleavage domain inactive (J. S. Chen et al. 2017; Casini et al. 2018). Recent biophysical studies using single-molecule FRET have revealed a highly dynamic conformation of SpCas9 where allosteric interactions between the PAM distal end and the HNH domain of the Cas9 enzyme renders a cleavage-impaired conformationally “closed” configuration upon encountering mismatches close to the 5' end of the sgRNA (Yang et al. 2018). To dissect if a greater stringency of target recognition could be achieved in case of FnCas9 by increasing the number of mismatches at the 5' end of the substrate, we selected 2 well-studied loci *EMX1* and *VEGFA3*, amplified their genomic off-targets with 2 and 3 mismatches at the non-seed region (PAM distal end), and interrogated the in vitro cleavage efficiency of FnCas9. Remarkably, FnCas9 was unable to cleave the substrate in the presence of 2 or 3 mismatches for both loci, suggesting that it is extremely specific in target recognition, particularly when the mismatches occur together in the PAM distal region (Fig1B). However, SpCas9 was able to cleave both of the mismatched substrates (Fig1B). Previous studies have highlighted the importance of defined mutations in the REC3 domain of highly specific engineered versions of Cas9 in determining target specificity by allosterically regulating the HNH domain from adopting a cleavage-competent form (J. S. Chen et al. 2017; Jones et al. 2021). However, the different engineered variants showed a similar binding affinity for their off-targets as wild-type SpCas9, even though they did not cleave these targets, suggesting that they probably remain bound to off-targets in a cleavage incompetent state (Jones et al. 2021; J. S. Chen et al. 2017). We asked if FnCas9 too shows similar properties and investigated the binding affinity for off-targets using MST. Strikingly, FnCas9 showed negligible to no binding affinity for

substrates having 2 mismatches where no cleavage was observed, suggesting that it either interacts extremely weakly or is evicted from the substrate following off-target interrogation (Fig1B). However, SpCas9 showed a strong binding affinity for both on- and off-targets, as reported previously (FigB) (J. S. Chen et al. 2017; Jones et al. 2021). Taken together, these results indicate that FnCas9 has a fundamentally distinct outcome of off-target recognition and binding as compared to SpCas9 and its engineered derivatives.

Intrigued by the high specificity of substrate recognition under in vitro conditions, we next investigated if FnCas9 can function in vivo as a genome-editing agent and how its genome-editing properties compared with that of SpCas9. For a bona-fide comparison between the 2 different Cas9 proteins with identical expression parameters, we generated constructs with either SpCas9 or FnCas9 and their corresponding sgRNAs from the same backbone. To select cells that received this plasmid by FACS sorting, we supplemented this construct with a T2A-eGFP sequence. To compare on-target and off-target editing, we selected 2 loci, *c-MYC* and *EMX1*, where genomic off-targets with 2 mismatches could be identified and 1 locus *HBB*, where genomic off-targets with 3 mismatches could be identified and therapeutically relevant. In the case of *EMX1*, these off-targets were earlier validated by an unbiased genome-wide GUIDE-Seq study (Tsai et al. 2015). We performed genome editing with SpCas9 and FnCas9 in HEK293T cells and a construct containing a scrambled sgRNA sequence was used as control. After sorting GFP⁺ cells from both SpCas9 and FnCas9 transfected cells, we isolated genomic DNA from these cells, amplified the target and off-target loci for each of the genes, and performed deep sequencing to quantify the number of insertions/deletions (indels). We observed that FnCas9 was able to cleave all of the 3 loci efficiently although with different efficiencies as compared to SpCas9. In the case of *c-MYC*, ~60% indels were seen, which dropped to ~22% for *EMX1* and ~19% for *HBB* (Fig1C). In contrast, SpCas9 showed greater than 50% indels for each of the 3 loci tested. Although SpCas9 cleaved more efficiently at each of the 3 loci, it showed modest to high cleavage at the off-targets for each locus. Among those with 2 mismatches, these included 1 off-target for *c-MYC* (~8%) and 2 off-targets for *EMX1* (~16% each). Even with 3 mismatches to the sgRNA, it showed ~2% cleavage for at least 1 off-target.

Strikingly, FnCas9 did not show any cleavage at any of the off-targets tested except *EMX1* off-target 2, where a very low ~1.15% of indels were detected (FigC). Although it might be argued that FnCas9 shows a generally lower efficiency of DNA cleavage due to which off-target cleavage events are diluted, the case of *c-MYC* where it exhibits comparable cleavage efficiency as SpCas9 yet maintains no cleavage at off-target loci rules out such an explanation. Collectively, we reported that FnCas9 did not show any off-target activity, a prerogative for therapeutic gene targeting. This part of work has been published in 2019 (Acharya et al. 2019).

Despite the high intrinsic specificity of FnCas9 action, low to modest cellular editing efficiency hinders its use as a robust genome editing tool for therapeutic applications. This motivates us to engineer this protein for enhanced activity without compromising on its function.

2. Development of an enhanced genome editing protein by structure-guided rational engineering

Although FnCas9 is evolutionarily divergent to SpCas9, it does share conserved functionality in terms of PAM recognition (Nishimasu et al. 2014; H. Hirano et al. 2016). PAM recognition by Cas9 triggers directional protospacer DNA unwinding, R-loop formation, and expansion which eventually reorients the HNH endonuclease domain to DNA cutting and concomitant RuvC activation leading to concerted DNA cleavage (Jiang and Doudna 2017). Recent mechanistic studies showed that the directional PAM-duplex DNA unwinding serves as the rate-limiting primary checkpoint of Cas9 action and a conformational switch exists between the Cas9 DNA binding and cleavage (Sternberg et al. 2014, 2015; Dagdas et al. 2017; Gong et al. 2018; Okafor et al. 2019). Moreover, the loss of nucleobase-specific interaction between the target DNA and Cas9 was reported to be rescued by base non-specific Cas9 interactions (H. Hirano et al. 2016; Nishimasu et al. 2018). Thus, we reasoned that stabilizing FnCas9 DNA duplex binding by introducing base non-specific interactions between PAM duplex and

the protein might improve FnCas9 nuclease activity without compromising its intrinsic specificity.

We engineered 50 different FnCas9 variants guided by its crystal structure bearing mostly single amino acid substitutions in the WED-PI domain to introduce novel PAM duplex DNA contacts. We took a substrate containing GGG PAM (where FnCas9 was shown to be least active (H. Hirano et al. 2016)) and performed target DNA cleavage experiments with the FnCas9 variants. We assayed for engineered FnCas9 (enFnCas9) variants with enhanced kinetic activity by *in vitro* cleavage screening assay (Fig2A). We selected a subset of 9 enFnCas9 variants which were ranked based on the number of amino acid substitutions and its position on the protein. Out of 9 variants, 5 enFnCas9 variants (en1, en2, en3, en4 and en15) harboring single amino acid substitution and 4 variants (en31, en34, en40 and en47) carrying a combination of single amino acid substitutions. en1, en15, en31 and en40 showed >80% substrate DNA cleaving activity within 0.5 min (Fig2A). Furthermore, en1, en15 and en31 showed around 2 fold higher cleavage activity w.r.t wild-type protein in the *in-vitro* time kinetics assay (Fig2B). Two independent groups had earlier reported that SpCas9 and its engineered altered-PAM variants with multiple amino acid substitutions create additional phosphate backbone interactions which synergistically induce unexpected displacement in the DNA backbone of the PAM duplex and facilitate these variants to recognize the non-canonical PAMs (S. Hirano et al. 2016; Anders, Bargsten, and Jinek 2016). To test if enFnCas9 variants show a similar relaxation in PAM recognition, we selected a subset of 5 enFnCas9 variants based on their enhanced activity at the non-canonical NGA PAM containing DNA substrates (Fig2C). Next, we performed an *in vitro* PAM discovery assay where a pooled library with a target DNA sequence (protospacer region) adjacent to a randomized 8 bp sequence ($4^8 = 65,536$ combinations in total) was cleaved with respective FnCas9 or enFnCas9-sgRNA complexes. The PAM depleted library was amplified and deep sequenced to comprehensively determine the PAM specificities of the variants. We observed that the enFnCas9 variants shared similar activity as FnCas9 at NGG PAMs but more flexible recognition in second and third nucleotide positions. Importantly, for all the variants tested (data shown for only en31 which has the highest NGA PAM activity), NGG was relaxed to NGR/NRG thereby expanding the scope of

enFnCas9 accessibility across the human genome (Fig2D). With this broadened PAM accessibility, enFnCas9 variants rank just below SpCas9-RY (Walton et al. 2020) and SpCas9-NG (Walton et al. 2020; Nishimasu et al. 2018) in their ability to target sites on the human genome when compared to other CRISPR systems reported to date (data not shown). The remarkable intrinsic specificity of FnCas9 to single-nucleotide mismatches in the target has applications both in disease diagnostics and disease correction. At the level of diagnostics, FnCas9 has recently been utilized for paper strip-based robust diagnostics of nucleic acid targets through the FnCas9 Editor Linked Uniform Detection Assay (FELUDA) (Azhar et al. 2021; Kumar et al. 2021) and Rapid Variant Assay (RAY) platforms (Kumar et al. 2021). These platforms utilize a direct FnCas9:DNA binding-based readout as opposed to collateral cleavage of DNA reporters as employed by Type V effectors (such as Cas12 or Cas14) (Kumar et al. 2021; J. S. Chen et al. 2018; Harrington et al. 2018) or RNA reporters cleavage after a targeted reverse transcription reaction (such as Cas13) (Gootenberg et al. 2017). We had earlier shown that placing an additional second mismatch in defined positions in the sgRNA (such as 2nd/6th PAM proximal or 16th/19th PAM distal positions) causes the enzyme to dissociate from its target and can be used as a readout for discriminating point mismatches such as disease-causing single nucleotide variants (SNVs) or Variants of Concern (VOCs) in SARS CoV2 infections (Azhar et al. 2021; Kumar et al. 2021). At a visual level, this can be used for diagnosing such mutants on a lateral flow device using FAM-labeled sgRNA (Fig3A). We reasoned that due to their broadened PAM accessibility, enFnCas9 variants (NRG/NGR) can now cover 78.27% of the reported Mendelian SNVs across the human genome (compared to 34.95% by FnCas9) thereby increasing the scope of detection to more disease-causing variants (data not shown). Expectedly, on a lateral flow strip, all enFnCas9 variants tested showed robust activity on a substrate carrying the non-canonical NGA PAM whereas FnCas9 did not show any signal (Fig3A). To determine if the kinetic changes in enFn variants might affect their diagnostic potential, we selected the enFnCas9 variant with the broadest activity at altered PAM sites (en31) and investigated if it was able to distinguish single mismatches in two targets with pathogenic mutations related to Sickle Cell Anemia (Fig3B) and the SARS CoV2 Alpha VOC signature (N501Y) (data not shown).

Remarkably en31 accurately distinguished both the target SNVs on a lateral flow device as compared to FnCas9 suggesting stronger enFnCas9:DNA interaction (Fig3B). We confirmed that the same specificity of SNV discrimination was also extended for an NGA PAM-containing substrate as well (data not shown). We reasoned that engineering FnCas9 by altering residues that interact with PAM in the substrate might have altered its binding affinity to DNA as evident from a stronger signal on a lateral flow device and more rapid enzymatic cleavage. To test this, we constructed recombinant Green Fluorescent Protein (GFP) tagged catalytically inactive versions of two of the variants (en1 and en15) and performed microscale thermophoresis (MST) to determine their DNA binding affinities on the same substrate (*VEGFA*). We found that these variants showed stronger DNA binding (en1 K_d 91.33 ± 29.8 nM, en15 K_d 49.16 ± 10.96 nM) as compared to FnCas9 (K_d 170 ± 31.53 nM) with en15 showing ~3.5 fold higher DNA binding affinity (Fig3C). Interestingly, in our previous study, we had seen that FnCas9 showed weaker binding to the same substrate as SpCas9 (3.02 fold) (Acharya et al. 2019). Thus, engineering improved enFnCas9:DNA binding affinity, reaching similar levels as SpCas9 but with superior specificity.

The safety of therapeutic genome editing is guided by off-target interrogation of CRISPR effectors. Although Cas12 and Cas14 have higher specificity than SpCas9, their therapeutic success relies on minimum ssDNA cleavage inside the cell such as those formed during replication, homology-directed repair, or transcription (J. S. Chen et al. 2018). On the contrary, enFnCas9 does not produce trans-cleavage products, and its high specificity both at the level of DNA interrogation and cleavage might be beneficial for safe nuclease-mediated genome editing. Although SpCas9 has shown robust gene editing capabilities across different genomic loci, the intrinsic non-specific nature of the protein has warranted the development of high fidelity versions for potential therapeutic editing (Tsai et al. 2015; Wienert et al. 2019). Interestingly, high fidelity SpCas9 proteins generally show lower editing efficiencies as compared to the wild-type protein. We selected two such proteins (SpCas9-HF1 and eSpCas9) due to their balanced activity and specificity as reported in literature (N. Kim et al. 2020). and compared their cellular editing rates (insertion/deletions) with enFnCas9 variants enFn1 and enFn15 in HEK293T cells. Among the two, enFn1 showed higher editing rates as compared to

high fidelity SpCas9 proteins at all 4 loci investigated and no detectable off-targets at any of the validated off-target sites (Fig3D). Expectedly, at all 4 loci, en1 had equal or higher editing efficiencies than the wildtype FnCas9 protein (Fig3D). These results showed that en1 achieves higher genome editing efficiency than high fidelity SpCas9 variants but retains similarly high on-target specificity. Finally, we investigated if the higher editing efficiency of the enFnCas9 variants can also lead to greater homology-directed repair (HDR) when presented with a double-stranded DNA (dsDNA) template. Here too, we observed higher HDR mediated knock-in of a donor (4.1 kb) at the *DCX* locus in HEK293T cells for both enFn1 and enFn15 as compared to SpCas9-HF1 and eSpCas9 (Fig3E).. Collectively enFn1 showed a higher rate of gene editing at all the target loci tested both for insertions/deletions as well HDR mediated knock-in highlighting its suitability as a highly potent genome-editing protein (Fig3D,E). Interestingly, the specificity of these variants appears to stem from the DNA interaction properties of FnCas9 independent of the engineered residues in the enzyme. Thus, we observed that even the strongest DNA binding enFnCas9 (en15) showed minimal editing at a GUIDE-Seq validated off-target with a single mismatch while substantially improving on-target activity compared to FnCas9 protein, whereas both eSpCas9 and Cas9-HF1 showed editing efficiencies comparable to the target site in good agreement with previous reports (Fig3F) (Kleinstiver et al. 2016; Casini et al. 2018). This shows that enFnCas9 variants possibly negotiate off-targets through a different mechanism than high fidelity SpCas9 proteins.

Taken together, enFnCas9 variants armed with superior activity and specificity provide an attractive platform for therapeutic correction of genetic diseases and disease diagnosis.

STATISTICAL ANALYSIS

Data are represented as means \pm SEM from independent experiments (mentioned in the legend if otherwise). MST analysis was done using NanoTemper analysis software. Determination of indel frequency from sequencing data was performed using

CRISPResso2 v2.0.29. DCX HDR data was analyzed using one-way Anova. A significance level of 0.05 was used throughout the study.

DISCUSSION

Cas9 belongs to the type II CRISPR system which has been efficiently exploited to edit a variety of genomes and has recently gained prominence for gene-therapy applications, predominantly for monogenic disorders like β -thalassemia and sickle cell anemia (Dever et al. 2016). Type II systems are classified in three subtypes IIA, IIB and IIC based on their evolutionarily divergent nature in terms of domain composition, overall size, primary sequence, gRNA feature and PAM specificity (Fonfara et al. 2014; Chylinski et al. 2014). SpCas9 protein which belongs to the type IIA system is one of the first described Cas9 proteins, gained prominence as a genome editing modality owing to its robust activity inside human cells but with a caution of paramount off-targeting effect. In our study, we described FnCas9 which belongs to the type IIB system as a safer alternative for genome editing and diagnostic platform owing to its specificity both *in vitro* and *in vivo* scenarios. The lower cellular activity and NGG PAM constraint which posed a challenge towards its applicability in disease correction and diagnostics was met by rational engineering. Thus, engineered FnCas9 variants hold a lot of promise in disease correction and diagnosis. Moreover, these variants open up lucrative avenues for developing double strand break free editing platforms (base editing and prime editing) where the nuclease system poses concerns for causing genotoxicity (Komor et al. 2016; Gaudelli et al. 2017; Anzalone, Koblan, and Liu 2020; Anzalone et al. 2019).

IMPACT OF THE RESEARCH

Our study has established a safe genome editing and diagnostic platform using FnCas9. Our report on the discovery of FnCas9 as a specific genome editing protein both *in vitro* and *in vivo* in 2019 has set up the platform for developing CRISPR diagnostics (CRISPRDx) platforms, FELUDA and RAY. FELUDA has already been deployed for detection of SARS Cov-2 in the wake of the Coronavirus pandemic by

TATAMD CHECK as CRISPR SARS Cov-2 Test. The engineered variants of FnCas9 render the FELUDA and RAY CRISPRDx platforms more robust and sensitive owing to their broad PAM specificities and stronger binding affinity with the target DNA. Furthermore, our group is proceeding towards pre-clinical studies with engineered FnCas9 for correcting the Sick cell anemia mutation with a goal to translate it to a therapeutic platform in the coming years.

REFERENCES

- Acharya, Sundaram, Souvik Maiti, and Debojyoti Chakraborty. 2020. "CRISPR-Cas9 for Therapy: The Challenges and Ways to Overcome Them." *Genome Engineering via CRISPR-Cas9 System*. <https://doi.org/10.1016/b978-0-12-818140-9.00009-x>.
- Acharya, Sundaram, Arpit Mishra, Deepanjan Paul, Asgar Hussain Ansari, Mohd Azhar, Manoj Kumar, Riya Rauthan, et al. 2019. "Cas9 Interrogates Genomic DNA with Very High Specificity and Can Be Used for Mammalian Genome Editing." *Proceedings of the National Academy of Sciences of the United States of America* 116 (42): 20959–68.
- Anders, Carolin, Katja Bargsten, and Martin Jinek. 2016. "Structural Plasticity of PAM Recognition by Engineered Variants of the RNA-Guided Endonuclease Cas9." *Molecular Cell* 61 (6): 895–902.
- Anzalone, Andrew V., Luke W. Koblan, and David R. Liu. 2020. "Genome Editing with CRISPR-Cas Nucleases, Base Editors, Transposases and Prime Editors." *Nature Biotechnology* 38 (7): 824–44.
- Anzalone, Andrew V., Peyton B. Randolph, Jessie R. Davis, Alexander A. Sousa, Luke W. Koblan, Jonathan M. Levy, Peter J. Chen, et al. 2019. "Search-and-Replace Genome Editing without Double-Strand Breaks or Donor DNA." *Nature* 576 (7785): 149–57.
- Azhar, Mohd, Rhythm Phutela, Manoj Kumar, Asgar Hussain Ansari, Riya Rauthan, Sneha Gulati, Namrata Sharma, et al. 2021. "Rapid and Accurate Nucleobase

Detection Using FnCas9 and Its Application in COVID-19 Diagnosis.” *Biosensors & Bioelectronics* 183 (July): 113207.

Casini, Antonio, Michele Olivieri, Gianluca Petris, Claudia Montagna, Giordano Reginato, Giulia Maule, Francesca Lorenzin, et al. 2018. “A Highly Specific SpCas9 Variant Is Identified by in Vivo Screening in Yeast.” *Nature Biotechnology* 36 (3): 265–71.

Chen, Fuqiang, Xiao Ding, Yongmei Feng, Timothy Seebeck, Yanfang Jiang, and Gregory D. Davis. 2017. “Targeted Activation of Diverse CRISPR-Cas Systems for Mammalian Genome Editing via Proximal CRISPR Targeting.” *Nature Communications* 8 (April): 14958.

Chen, Janice S., Yavuz S. Dagdas, Benjamin P. Kleinstiver, Moira M. Welch, Alexander A. Sousa, Lucas B. Harrington, Samuel H. Sternberg, J. Keith Joung, Ahmet Yildiz, and Jennifer A. Doudna. 2017. “Enhanced Proofreading Governs CRISPR-Cas9 Targeting Accuracy.” *Nature* 550 (7676): 407–10.

Chen, Janice S., Enbo Ma, Lucas B. Harrington, Maria Da Costa, Xinran Tian, Joel M. Palefsky, and Jennifer A. Doudna. 2018. “CRISPR-Cas12a Target Binding Unleashes Indiscriminate Single-Stranded DNase Activity.” *Science* 360 (6387): 436–39.

Chylinski, Krzysztof, Kira S. Makarova, Emmanuelle Charpentier, and Eugene V. Koonin. 2014. “Classification and Evolution of Type II CRISPR-Cas Systems.” *Nucleic Acids Research*. <https://doi.org/10.1093/nar/gku241>.

Cong, Le, F. Ann Ran, David Cox, Shuailiang Lin, Robert Barretto, Naomi Habib, Patrick D. Hsu, et al. 2013. “Multiplex Genome Engineering Using CRISPR/Cas Systems.” *Science* 339 (6121): 819–23.

Dagdas, Yavuz S., Janice S. Chen, Samuel H. Sternberg, Jennifer A. Doudna, and Ahmet Yildiz. 2017. “A Conformational Checkpoint between DNA Binding and Cleavage by CRISPR-Cas9.” *Science Advances* 3 (8): eaao0027.

Dever, Daniel P., Rasmus O. Bak, Andreas Reinisch, Joab Camarena, Gabriel Washington, Carmencita E. Nicolas, Mara Pavel-Dinu, et al. 2016. “CRISPR/Cas9 β -Globin Gene Targeting in Human Haematopoietic Stem Cells.” *Nature* 539 (7629): 384–89.

- Esvelt, Kevin M., Prashant Mali, Jonathan L. Braff, Mark Moosburner, Stephanie J. Yang, and George M. Church. 2013. "Orthogonal Cas9 Proteins for RNA-Guided Gene Regulation and Editing." *Nature Methods* 10 (11): 1116–21.
- Fonfara, Ines, Anaïs Le Rhun, Krzysztof Chylinski, Kira S. Makarova, Anne-Laure Lécroivain, Janek Bzdrenga, Eugene V. Koonin, and Emmanuelle Charpentier. 2014. "Phylogeny of Cas9 Determines Functional Exchangeability of Dual-RNA and Cas9 among Orthologous Type II CRISPR-Cas Systems." *Nucleic Acids Research* 42 (4): 2577–90.
- Gaudelli, Nicole M., Alexis C. Komor, Holly A. Rees, Michael S. Packer, Ahmed H. Badran, David I. Bryson, and David R. Liu. 2017. "Programmable Base Editing of A•T to G•C in Genomic DNA without DNA Cleavage." *Nature* 551 (7681): 464–71.
- Gong, Shanzhong, Helen Hong Yu, Kenneth A. Johnson, and David W. Taylor. 2018. "DNA Unwinding Is the Primary Determinant of CRISPR-Cas9 Activity." *Cell Reports* 22 (2): 359–71.
- Gootenberg, Jonathan S., Omar O. Abudayyeh, Jeong Wook Lee, Patrick Essletzbichler, Aaron J. Dy, Julia Joung, Vanessa Verdine, et al. 2017. "Nucleic Acid Detection with CRISPR-Cas13a/C2c2." *Science* 356 (6336): 438–42.
- Harrington, Lucas B., David Burstein, Janice S. Chen, David Paez-Espino, Enbo Ma, Isaac P. Witte, Joshua C. Cofsky, Nikos C. Kyrpides, Jillian F. Banfield, and Jennifer A. Doudna. 2018. "Programmed DNA Destruction by Miniature CRISPR-Cas14 Enzymes." *Science* 362 (6416): 839–42.
- Hirano, Hisato, Jonathan S. Gootenberg, Takuro Horii, Omar O. Abudayyeh, Mika Kimura, Patrick D. Hsu, Takanori Nakane, et al. 2016. "Structure and Engineering of Francisella Novicida Cas9." *Cell* 164 (5): 950–61.
- Hirano, Seiichi, Hiroshi Nishimasu, Ryuichiro Ishitani, and Osamu Nureki. 2016. "Structural Basis for the Altered PAM Specificities of Engineered CRISPR-Cas9." *Molecular Cell* 61 (6): 886–94.
- Jiang, Fuguo, and Jennifer A. Doudna. 2017. "CRISPR–Cas9 Structures and Mechanisms." *Annual Review of Biophysics*. <https://doi.org/10.1146/annurev-biophys-062215-010822>.
- Jones, Stephen K., Jr, John A. Hawkins, Nicole V. Johnson, Cheulhee Jung, Kuang Hu,

- James R. Rybarski, Janice S. Chen, Jennifer A. Doudna, William H. Press, and Ilya J. Finkelstein. 2021. "Massively Parallel Kinetic Profiling of Natural and Engineered CRISPR Nucleases." *Nature Biotechnology* 39 (1): 84–93.
- Kim, Eunji, Taeyoung Koo, Sung Wook Park, Daesik Kim, Kyoungmi Kim, Hee-Yeon Cho, Dong Woo Song, et al. 2017. "In Vivo Genome Editing with a Small Cas9 Orthologue Derived from *Campylobacter Jejuni*." *Nature Communications* 8 (February): 14500.
- Kim, Nahye, Hui Kwon Kim, Sungtae Lee, Jung Hwa Seo, Jae Woo Choi, Jinman Park, Seonwoo Min, Sungroh Yoon, Sung-Rae Cho, and Hyongbum Henry Kim. 2020. "Prediction of the Sequence-Specific Cleavage Activity of Cas9 Variants." *Nature Biotechnology* 38 (11): 1328–36.
- Kleinstiver, Benjamin P., Vikram Pattanayak, Michelle S. Prew, Shengdar Q. Tsai, Nhu T. Nguyen, Zongli Zheng, and J. Keith Joung. 2016. "High-Fidelity CRISPR-Cas9 Nucleases with No Detectable Genome-Wide off-Target Effects." *Nature* 529 (7587): 490–95.
- Komor, Alexis C., Yongjoo B. Kim, Michael S. Packer, John A. Zuris, and David R. Liu. 2016. "Programmable Editing of a Target Base in Genomic DNA without Double-Stranded DNA Cleavage." *Nature* 533 (7603): 420–24.
- Kumar, Manoj, Sneha Gulati, Asgar H. Ansari, Rhythm Phutela, Sundaram Acharya, Mohd Azhar, Jayaram Murthy, et al. 2021. "FnCas9-Based CRISPR Diagnostic for Rapid and Accurate Detection of Major SARS-CoV-2 Variants on a Paper Strip." *eLife* 10 (June). <https://doi.org/10.7554/eLife.67130>.
- Mali, Prashant, Luhan Yang, Kevin M. Esvelt, John Aach, Marc Guell, James E. DiCarlo, Julie E. Norville, and George M. Church. 2013. "RNA-Guided Human Genome Engineering via Cas9." *Science* 339 (6121): 823–26.
- Nishimasu, Hiroshi, F. Ann Ran, Patrick D. Hsu, Silvana Konermann, Soraya I. Shehata, Naoshi Dohmae, Ryuichiro Ishitani, Feng Zhang, and Osamu Nureki. 2014. "Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA." *Cell* 156 (5): 935–49.
- Nishimasu, Hiroshi, Xi Shi, Soh Ishiguro, Linyi Gao, Seiichi Hirano, Sae Okazaki, Taichi Noda, et al. 2018. "Engineered CRISPR-Cas9 Nuclease with Expanded Targeting

Space.” *Science* 361 (6408): 1259–62.

- Okafor, Ikenna C., Digvijay Singh, Yanbo Wang, Minhee Jung, Haobo Wang, John Mallon, Scott Bailey, Jungjoon K. Lee, and Taekjip Ha. 2019. “Single Molecule Analysis of Effects of Non-Canonical Guide RNAs and Specificity-Enhancing Mutations on Cas9-Induced DNA Unwinding.” *Nucleic Acids Research* 47 (22): 11880–88.
- Ran, F. Ann, Le Cong, Winston X. Yan, David A. Scott, Jonathan S. Gootenberg, Andrea J. Kriz, Bernd Zetsche, et al. 2015. “In Vivo Genome Editing Using *Staphylococcus Aureus* Cas9.” *Nature* 520 (7546): 186–91.
- Sternberg, Samuel H., Benjamin LaFrance, Matias Kaplan, and Jennifer A. Doudna. 2015. “Conformational Control of DNA Target Cleavage by CRISPR-Cas9.” *Nature* 527 (7576): 110–13.
- Sternberg, Samuel H., Sy Redding, Martin Jinek, Eric C. Greene, and Jennifer A. Doudna. 2014. “DNA Interrogation by the CRISPR RNA-Guided Endonuclease Cas9.” *Nature* 507 (7490): 62–67.
- Tsai, Shengdar Q., Zongli Zheng, Nhu T. Nguyen, Matthew Liebers, Ved V. Topkar, Vishal Thapar, Nicolas Wyvekens, et al. 2015. “GUIDE-Seq Enables Genome-Wide Profiling of off-Target Cleavage by CRISPR-Cas Nucleases.” *Nature Biotechnology* 33 (2): 187–97.
- Walton, Russell T., Kathleen A. Christie, Madelynn N. Whittaker, and Benjamin P. Kleinstiver. 2020. “Unconstrained Genome Targeting with near-PAMless Engineered CRISPR-Cas9 Variants.” *Science* 368 (6488): 290–96.
- Wienert, Beeke, Stacia K. Wyman, Christopher D. Richardson, Charles D. Yeh, Pinar Akcakaya, Michelle J. Porritt, Michaela Morlock, et al. 2019. “Unbiased Detection of CRISPR off-Targets in Vivo Using DISCOVER-Seq.” *Science* 364 (6437): 286–89.
- Yang, Mengyi, Sijia Peng, Ruirui Sun, Jingdi Lin, Nan Wang, and Chunlai Chen. 2018. “The Conformational Dynamics of Cas9 Governing DNA Cleavage Are Revealed by Single-Molecule FRET.” *Cell Reports* 22 (2): 372–82.

A)

B)

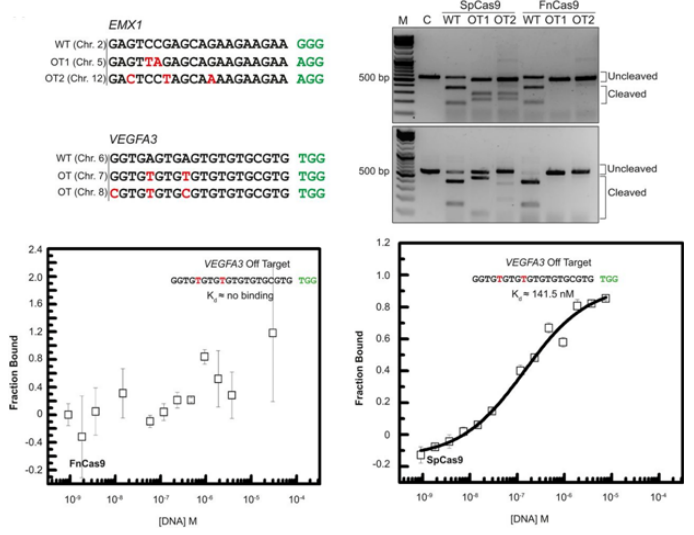


Figure 2 consists of three bar charts showing the percentage of indels for SpCas9 (red) and FnCas9 (cyan) across on-target and off-target sites for three different sgRNAs: *c-MYC*, *EMX1*, and *HBB*.

Left Chart: *c-MYC* sgRNA

Category	Target	SpCas9 (%)	FnCas9 (%)
On target	T	~62	~62
	C	~68	~3
	T	~6	~0
	C	~0	~0
Off target 1 (2MM)	T	~8	~0
	C	~0	~0
	T	~0	~0
	C	~0	~0
Off target 2 (2MM)	T	~0	~0
	C	~0	~0
	T	~0	~0
	C	~0	~0

Middle Chart: *EMX1* sgRNA

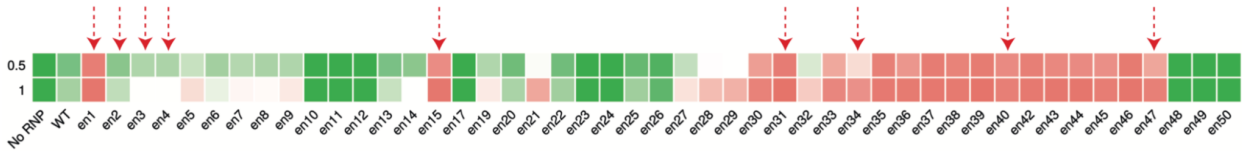
Category	Target	SpCas9 (%)	FnCas9 (%)
On target	T	~22	~22
	C	~50	~0
	T	~0	~0
	C	~0	~0
Off target 1 (2MM)	T	~2	~2
	C	~16	~0
	T	~2	~0
	C	~0	~0
Off target 2 (2MM)	T	~2	~2
	C	~17	~0
	T	~0	~0
	C	~0	~0

Right Chart: *HBB* sgRNA

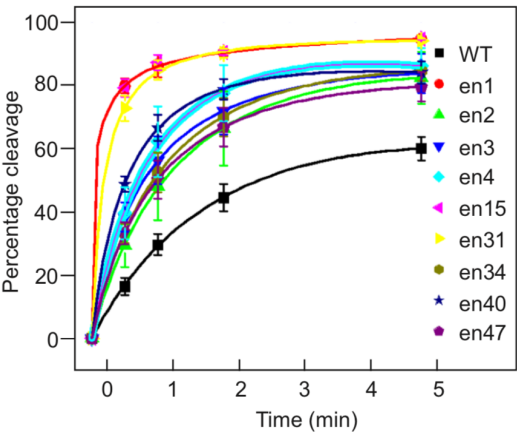
Category	Target	SpCas9 (%)	FnCas9 (%)
On target	T	~20	~20
	C	~59	~0
	T	~0	~0
	C	~0	~0
Off target 1 (3MM)	T	~0	~0
	C	~0	~0
	T	~0	~0
	C	~0	~0
Off target 2 (3MM)	T	~0	~0
	C	~0	~0
	T	~0	~0
	C	~0	~0

Figure 2

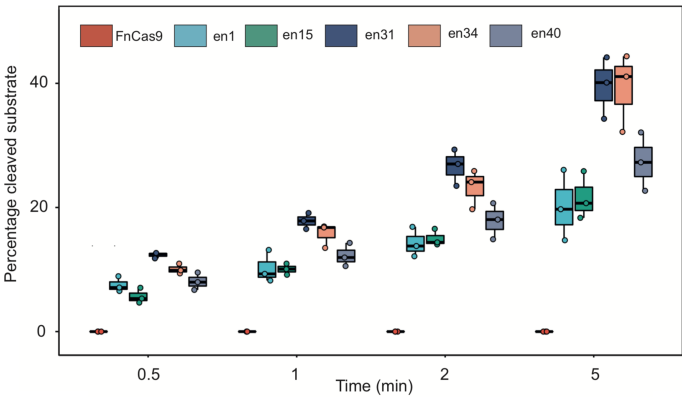
A)



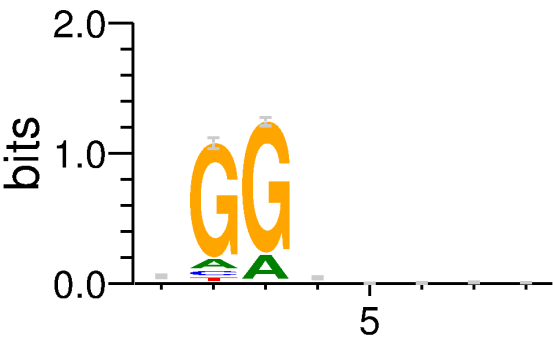
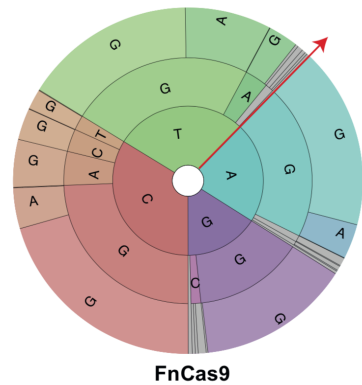
B)



C)



D)



E)

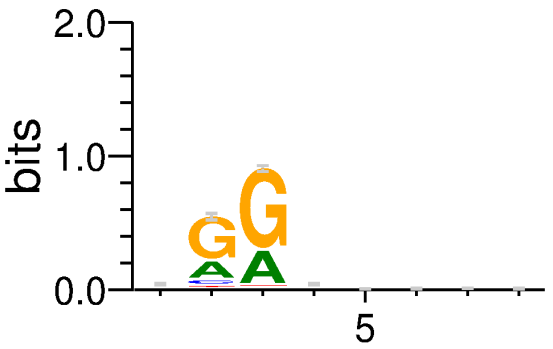
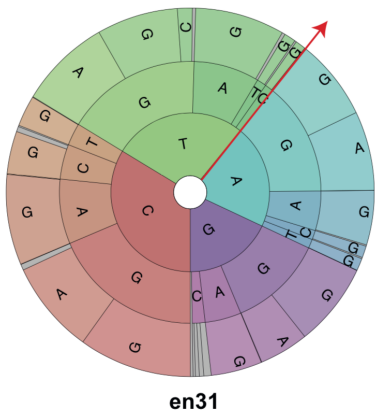
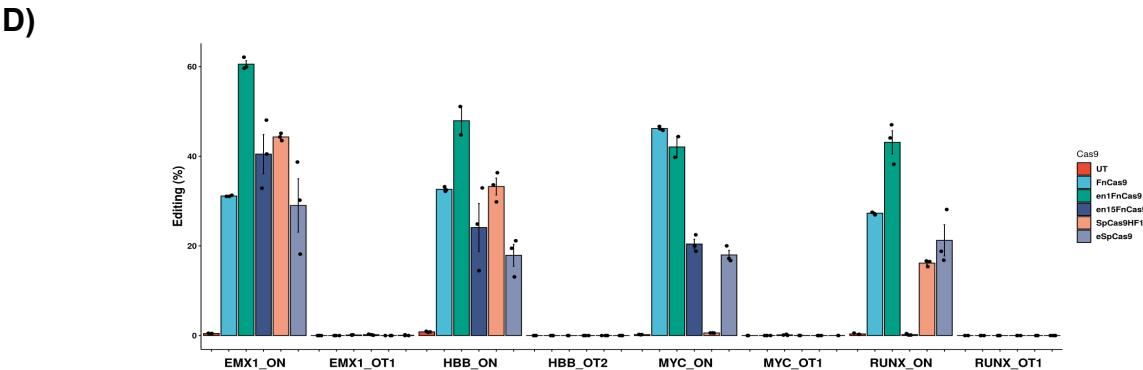
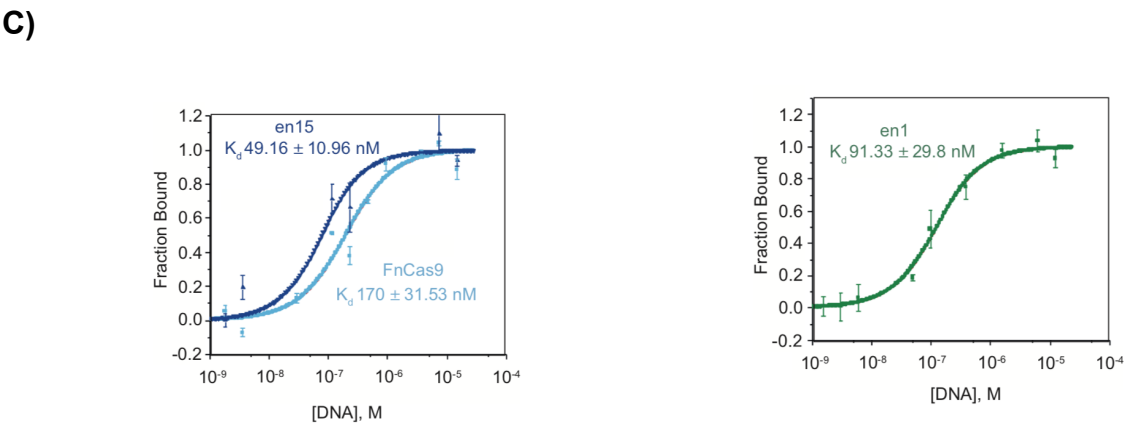
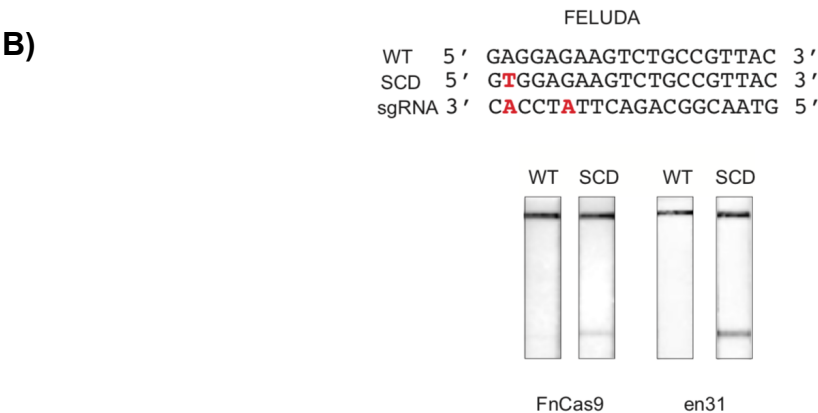
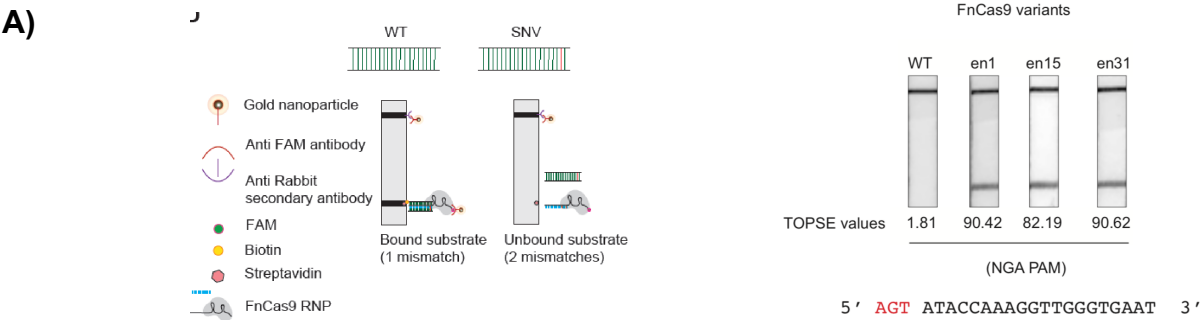
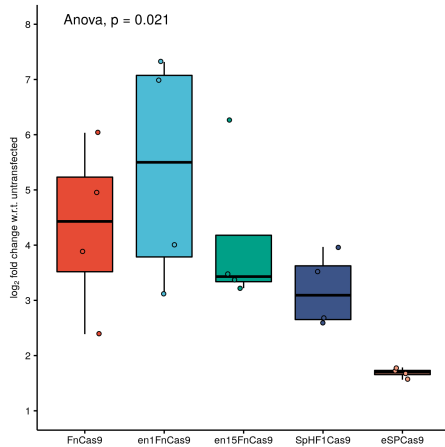


Figure 3



E)



F)

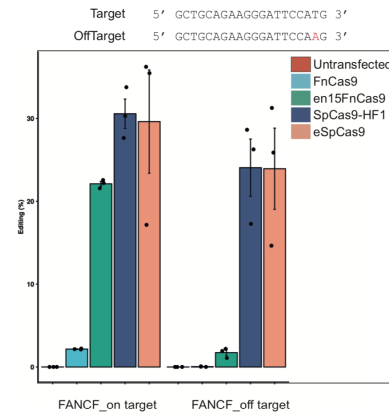


Figure1

Mismatches (MM) in target abrogates FnCas9 cleavage activity. A) Mismatches at different positions along the substrate alter FnCas9 cleavage outcomes while retaining SpCas9 activity. Representative in vitro cleavage outcome of HBB substrate (purified PCR product, 100 ng) showing mismatches at a single position in every lane (indicated in red, other bases remaining unchanged) interrogated with FnCas9 (*Upper*) or SpCas9 (*Lower*) in the form of RNP complexes (250 nM) is shown, PAM is indicated in green. The top band represents an uncleaved target while the 2 bottom bands represent cleaved products. Dotted box represents positions showing minimum tolerance to mismatches for FnCas9. Quantification for each reaction is shown below the gel images. Error bars represent SD (3 independent experiments)

FnCas9 does not bind to mismatched substrates. B) FnCas9 does not tolerate 2 or more mismatches in the substrate. In vitro cleavage outcomes of *EMX1* (*Upper*) and *VEGFA3* (*Lower*) targets interrogated by wild-type FnCas9 and SpCas9 are shown (*Right*). (*Left*) The sequences with 2 or more mismatches (indicated as red bases) are presented. PAM is indicated in green. Data representative of 3 independent experiments. MST results showing binding outcomes of dFnCas9-GFP (*Left*) and dSpCas9-GFP (*Right*) to *VEGFA3* off-target substrate expressed as fraction bound (y axis) with respect to varying concentrations of purified DNA substrate (x axis). The substrate sequences are indicated in the box with the PAM shown in green. Error bars represent SEM (2 independent experiments).

FnCas9 performs genome editing with very high specificity. C) Indel events

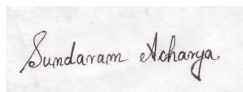
(expressed as percentage) obtained by amplicon sequencing upon SpCas9 or FnCas9 targeting *c-MYC*, *EMX1* and *HBB* sites along with corresponding off-target sites in HEK293T cells. Individual data points from independent experiments are shown.

Figure2

Characterization of engineered FnCas9 variants with enhanced kinetic activity and altered PAM specificities. A) Heat map showing the *in vitro* cleavage screening assay on GGG PAM containing DNA substrate for 0.5 and 1 min. B) *In vitro* cleavage kinetics assay showing percent cleavage (y-axis) as a function of time (x-axis). Error bars represent SD of three independent experiments. C) Box plot showing the in-vitro cleavage kinetics assay on NGA PAM containing DNA substrate. Percent cleavage was plotted on y-axis and the time points on x-axis. Error bars represent SD of three independent experiments. D) PAM wheel and sequence logo of FnCas9 showing predominant preference for NGG PAM. E) PAM wheel and sequence logo of en31 showing canonical NGG PAM preference is being relaxed to NGR/NRG.

Figure3

SNV detection and cellular genome editing by enFnCas9. A) Scheme of SNV detection by FELUDA and RAY (left). Detection of NGA PAM substrate on paper strip assay (right). B) Discriminating SCD by FELUDA where en31 showed more sensitive detection compared to FnCas9. C) MST assay showing improvement of binding affinity in en1 and en15 compared to FnCas9. Comparative binding affinity plot between en15 and FnCas9 (left). Binding affinity plot for en1 (right). D) Indel events (expressed as editing percentage) obtained by amplicon sequencing for Cas9s targeting respective loci along with corresponding off-target sites in HEK293T cells. Individual data points from independent experiments are shown. E) HDR assay targeting *DCX* locus showing higher efficiency of donor DNA template integration by en1. F) Indel events (expressed as editing percentage) obtained by amplicon sequencing for Cas9s targeting *FANCF* locus and the GUIDE-Seq validated off-target locus within single mismatch.



Sundaram Acharya
Senior Research Fellow-ICMR
CSIR-IGIB, New Delhi 110025

