International Journal of Biological Macromolecules 147 (2020) 376–384

Contents lists available at ScienceDirect

International Journal of Biological Macromolecules

journal homepage: http://www.elsevier.com/locate/ijbiomac

A novel thermal Cas12b from a hot spring bacterium with high target mismatch tolerance and robust DNA cleavage efficiency

Ye Tian [a,c], Rui-Rui Liu [a,c], Wen-Dong Xian [a,c], Mengjie Xiong [a,c], Min Xiao [a,c], Wen-Jun Li [a,b,c,*]

[a] State Key Laboratory of Biocontrol, Guangdong Provincial Key Laboratory of Plant Resources, School of Life Sciences, Sun Yat-Sen University, 510275 Guangzhou, PR China
[b] CAS Key Laboratory of Biogeography and Bioresource in Arid Land, Xinjiang Institute of Ecology and Geography, Urumqi 830011, PR China
[c] Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), Sun Yat-Sen University, 510275 Guangzhou, PR China

ARTICLE INFO

Article history:
Received 6 November 2019
Received in revised form 30 December 2019
Accepted 7 January 2020
Available online 9 January 2020

Keywords:
Thermophiles
CRISPR-Cas12b
Dynamic analysis

ABSTRACT

Clustered regularly interspaced short palindromic repeats (CRISPR)-associated proteins (Cas), such as Cas9 and Cpf1, are RNA-guided endonucleases that target and degrade nucleic acids, providing powerful genomic editing and molecular diagnostic tools. Cas12b enzymes are distinct effectors; however, their features and catalytic boundaries require further characterization. We identified BrCas12b from the thermophile bacterium *Brevibacillus* sp. SYSU G02855 as a novel ortholog of cas12b. Biochemical analyses revealed that BrCas12b is a dual-RNA-guided endonuclease with higher optimum reaction temperature than that of other reported members of Cas12b. The seed sequence of BrCas12b is only 4 nt in length, indicating that it has greater target mismatch tolerance than that of previously reported Cas effectors; however, it contains a compensatory effect at the position of the cleavage site. Using fluorescence-based detection method to evaluate target cleavage efficiency, we showed that BrCas12b has robust enzymatic cleavage activity (K_{cat}/K_m (s^{-1} M^{-1}) = 8.80×10^{11}), which is significantly higher than that of AacCas12b (K_{cat}/K_m (s^{-1} M^{-1}) = 7.56×10^{8}) from *Alicyclobacillus acidoterrestris*. The results increase our understanding of the catalytic mechanism of Cas12b family members and suggest that BrCas12b might be useful in the application of genomic editing and molecular diagnosis.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

Almost all archaea and many bacteria possess an adaptive immunological system that resists invasion of exogenous nucleic acid, such as virus and plasmids [1–3]. The genetic interference pathway of this immunological system consists of clustered regularly interspaced short palindromic repeats (CRISPR) and CRISPR-associated proteins (Cas) [4–6]. A CRISPR array consists of small variable spacers, separated by repeats. The spacers, as the immunological memory, not evolutionarily conserved, generally match and originate from mobile genetic elements, such as viruses and plasmids [5,7]. Typically, small fragments of the invading genetic elements are directionally integrated into CRISPR arrays, forming new spacers along a determined direction. Primary transcripts of the CRISPR array are processed to generate CRISPR RNA (crRNA), which binds to an effector (Cas protein) and guides the effector to match the target, achieving target degradation based on its binary complementary sequence [8,9]. Beyond these common characteristics, recent studies have uncovered new functional features of the CRISPR-Cas system, such as collateral nonspecific degradation of nucleic acids [10,11] and immunoreaction enhanced by secondary messengers [12,13].

Phylogenetically, CRISPR-Cas systems are of great diversity and can be classified into multiple types and subtypes, which in turn are clustered into two classes. The various types in Class 2 originated from the corresponding types in Class 1. Translocation events actuated multiple substitutions of the Class 1 proteins to form the various types of the Class 2 system [14]. This viewpoint is consistent with the shared kernel characteristics of the CRISPR-Cas system. Significantly, the difference between Class 1 and Class 2 is the effector complexes, which possess different numbers of Cas proteins. The Class 1 system achieves target destruction *via* a multi-subunit effector, whereas the Class 2 system contains a single, larger effector to perform target interference [14]. Class 2 effectors are hugely important in genome editing and molecular diagnosis [15,16].

Enzymes from the Class 2 CRISPR-Cas system, such as Cas9 and Cpf1, have been harnessed as versatile and reprogrammable molecular tools for genome editing, transcriptional regulation, and clinical diagnosis [17–23]. However, only two members of the Class 2 effector proteins are widely used as gene editing tools: Cas9, the forerunner of the CRISPR-Cas application, a dual-RNA-guided endonuclease that cleaves

* Corresponding author at: State Key Laboratory of Biocontrol, Guangdong Provincial Key Laboratory of Plant Resources and Southern Marine Science and Engineering Guangdong Laboratory (Zhuhai), School of Life Sciences, Sun Yat-Sen University, 510275 Guangzhou, PR China.
E-mail address: liwenjun3@mail.sysu.edu.cn (W.-J. Li).

https://doi.org/10.1016/j.ijbiomac.2020.01.079
0141-8130/© 2020 Elsevier B.V. All rights reserved.

target DNA and generates blunt-end products *via* the HNH and RuvC nuclease domains [24]; and Cpf1, a single crRNA-guided Cas nuclease that recognizes dsDNA substrate to produce 5′ staggered ends and possesses inducible indiscriminate ssDNase activity [10,11,25]. Recently, a new member of the Class 2 system, CRISPR-C2c1/Cas12b, has been discovered. Distinct from Cas9, the enzyme Cas12b contains only the RuvC domain, recognizes the thymine nucleotide-rich protospacer adjacent motif (PAM), and generates a product containing the staggered end [26,27]. Unlike the Cpf1 enzyme, Cas12b requires dual RNAs as the guide to recognize target DNA. Most notably, Cas12b has minimal off-target effects, providing a safer alternative for genomic editing and diagnostic applications [28]. However, details of the substrate specificity and reaction dynamics of the CRISPR-Cas12b system are unknown. The discovery of more members of Cas12b will contribute to a more comprehensive understanding of the characteristics of the CRISPR-Cas12b system.

Here, we report a novel Cas12b homolog obtained from *Brevibacillus* sp. SYSU G02855 (GenBank accession number: MN744355), a new species isolated from a hot spring in Tibet, China by high-temperature directional screening. Amino acid sequence analyses show that the BrCas12b protein shares a 34% sequence identity with the previously reported AacCas12b protein from *Alicyclobacillus acidoterrestris* [26]. BrCas12b possesses a number of features, such as higher optimum reaction temperature, greater target mismatch tolerance, and robust DNA cleavage efficiency, which make this novel enzyme a model to further our understanding of the workings of the Cas12b system.

2. Materials and methods

2.1. BrCas12b identification, expression, and purification

BrCas12b was identified *via* genomic sequencing data obtained from *Brevibacillus* sp. SYSU G02855, which was isolated from a hot spring. In the CRISPR-Cas system identification pipeline, the profile of Cas1 was used as the query in a basic local alignment search tool (BLAST) search against the BrCas12b genome, with an E-value cut-off of 0.01 and low-complexity filtering disabled [14]. The *cas1* gene was used as an anchor for scanning and identifying possible loci within the 20-kb region around the cas1 gene; this region was extracted and translated using GeneMarkS [29]. Predicted proteins were annotated using the National Center for Biotechnology Information non-redundant protein database with the PSI-Blast program. Protein sequence alignments were conducted using the MUSCLE program [30]. The FastTree program was used to perform a phylogenetic analysis [31]. The secondary structure of the protein was predicted using Jpred4 [32].

The full-length *BrCas12b* gene was amplified by PCR using Q5 High-Fidelity DNA Polymerase (NEB) and the primers 5′-taagaaggagatatacatATG(*Nde*I)CCTGTAAGAAGCTTCAAAGTAAA-3′ (forward) and 5′-ccctcgggtagggcaactagt(*Spe*I)gcatctcccgtgatgcaCTGAAG TTGCTCTCGCAGCTT-3′ (reverse). *Brevibacillus* sp. SYSU G02855 genomic DNA was used as the template. The PCR product was introduced into a pTXB plasmid *via* the *Nde*I/*Spe*I site to construct a fusion protein expression cassette between BrC2c1 and a self-cleaving affinity tag consisting of an Mxe intein portion and the chitin-binding fragment. The constructs were controlled by sequencing.

Recombinant BrCas12b protein was generated in the *Escherichia coli* BL21 (DE3) (Takara). The strain harboring the expression vector was cultured in lysogeny broth medium containing 100 μg/mL ampicillin at 37 °C. When the OD_{600} reached 0.5–0.6, IPTG was added to a final concentration of 0.5 mM. The cells were then incubated for 4 h at 28 °C. The cells were collected by centrifugation at 10,000 ×*g* for 5 min at 4 °C, and the pellet was stored at −80 °C.

Cells expressing the BrCas12b protein were resuspended in lysis buffer (20 mM Tris/HCl pH 8.0, 2 mM Na_2-EDTA, 1.2% Triton, and 1 mM PMSF) and incubated at 37 °C for 30 min. The treated cells were then crushed by ultrasonication, and the cell fragments were fractionated using 5 M NaCl, 10% (v/v) glycerol, 0.1 mmol/L DTA, and 1 mmol/L DTT. The elution products, including BrCas12b, were collected and concentrated to 5 mg/mL, and stored at −80 °C.

2.2. Bacterial RNA sequencing

Twenty milliliters of *Brevibacillus* sp. SYSU G02855 was harvested at the log phase of growth, centrifuged, and resuspended in 1 mL of lysis buffer (20 mM Tris, pH 8.0, 2 mM Na_2-EDTA, 1.2%Triton, 20 mg/mL lysozyme). RNA was then extracted, using an RNAprep Pure Cell/Bacteria Kit (TIANGEN). The purified RNA was then treated with DNase I (TAKARA), and the 3′ end was dephosphorylated using T4 polynucleotide kinase. Ribosome RNA was eliminated using the bacterial Ribo-Zero rRNA Removal Kit (Illumina). RNA libraries were generated from the rRNA-depleted RNA using the NEBNext Small RNA Library Prep Set for Illumina (New England Biolabs) and size selected using PippinPrep (Sage Science). The prepared libraries were then sequenced on a MiSeq system (Illumina).

2.3. RNA sequencing analysis

Producing reads were identified on the basis of their associated barcodes and aligned to the appropriate RefSeq reference genome using the Burrows-Wheeler Aligner (BWA) [33]. Paired-end alignments were used to extract entire transcript sequences using Galaxy (https://usegalaxy.org), and these sequences were analyzed using Geneious 8.1.8. For the remaining loci, putative tracrRNA (trans-acting CRISPR RNA) were identified based on the presence of an anti-direct repeat sequence. Anti-direct repeats were identified using Geneious 8.1.8. The transcriptional start sites of crRNA and tracrRNA were predicted using the BPROM (www.softberry.com) and ARNOLD (rna.lgmors.upsud.fr/toolbox/Arnold/) programs, respectively. RNA folding was analyzed using NUPACK software.

2.4. In vitro transcription and purification of RNA

In vitro transcription was carried out using a Thermo Scientific Transcript Aid T7 High Yield Transcription Kit. The target template, carrying a T7 promoter sequence (Supplementary Table S1), was amplified using thermal cycling. The crRNA and tracrRNA products were gel-purified and stored at −70 °C.

2.5. PAM screen

To analyze the requirement of the PAM, we generated random PAM libraries consisting of oligonucleotides constructed by Integrated DNA Technologies (IDT) and integrated into pMD19-T plasmids (TAKARA). The synthesized oligonucleotides (Supplementary Table S2) contain a spacer 1 sequence upstream that was modified by addition of a random 7 nt sequence. Before integration into the plasmid, the inserting oligonucleotides were made double stranded by annealing in a thermocycler at 95 °C for 5 min, with gradual cooling to 4 °C. The dsDNAs were inserted into pMD19-T vectors using DNA ligase. Top10 *E. coli* cells were transformed using the recombinant vector, after which the cells were harvested, and the plasmid DNA was isolated using a GenStar StarPrep Plasmid Miniprep Kit. The purified plasmids were digested by BrCas12b-dual RNA complexes at 62 °C for 90 min *in vitro*. Next, the products of digestion were loaded onto a 1.0% agarose gel and separated by electrophoresis. Bands of supercoiled plasmid DNA were excised from the gel and purified. The plasmid library was sequenced using an Illumina MiSeq with 150 single-end cycles. We performed the sequencing of the plasmid libraries before or after the treatment, and used the overlap of two sequences to identify the required PAM. The high-frequency patterns were identified by statistical analysis and demonstrated by electrophoresis on 1.0% agarose gel containing SYBR safe DNA Gel Stain (Thermo Fisher Scientific).

2.6. In vitro cleavage assays

Plasmid templates were generated by cloning a PAM-spacer 1 component into a pMD19-T plasmid, followed by purification using a StarPrep Plasmid Miniprep Kit (GenStar). Linear DNA templates were produced by PCR amplification of the corresponding plasmids and purified using a StarPrep PCR & DNA Fragment Purification Kit. The dual-RNA was pre-annealed by heating at 95 °C for 5 min followed by gradual cooling to room temperature. The *in vitro* DNA cleavage assays were carried out in a cleavage buffer consisting of 100 mM NaCl, 50 mM Tris-HCl, 10 mM $MgCl_2$, and 100 μg/mL BSA. The reaction mixture contained 8 nM Cas12b protein, 20 nM crRNA-tracrRNA complex, and 1 μg of template DNA in cleavage buffer, incubated at 62 °C for 1 h. This temperature was used for all reactions unless otherwise noted. The enzyme reactions were carried out in a thermocycler with the following procedure: 25 °C for 2 min, 62 °C for 60 min, then hold at 4 °C. To determine the optimal temperature for BrCas12b, a cleavage assay was conducted over a wide range of temperatures (30–80 °C) for 1 h in cleavage buffer. For the metal-dependent cleavage assay, cleavage buffer containing 2 or 10 mM $CaCl_2$, $NiSO_4$, $BaCl_2$, $MnCl_2$, $CoCl_2$, or $PbCl_2$ was used in place of buffer containing $MgCl_2$. The template, containing a spacer 1 sequence and a TTT-patterned PAM, was incubated in cleavage buffer with 8 nM BrCas12b and 20 nM pre-annealed dual-RNA at 62 °C for 1 h. The product was analyzed by electrophoresis on a 1.5% agarose gel containing SYBR Safe DNA Gel Stain.

2.7. Cleavage efficiency assay based on fluorometric measurements

The fluorometric measurement method was performed as described previously with modifications [34]. The ssDNA probes were labeled with 6-FAM at the 5′ site and with BHQ1 at the 3′ site (Supplementary Table S1). The dsDNA beacons were generated from ssDNA probes annealed in a thermocycler at 95 °C for 5 min, with gradual cooling to 4 °C. The fluorometric experiments were performed using beacons in cleavage buffer with the BrCas12b complex at 62 °C. The fluorescence emission was measured at 520 nm with excitation at 498 nm. The initial fluorescence of the samples was subtracted from the final fluorescence results to allow for comparisons among different conditions. The fluorescence reading obtained under background conditions, in the absence of the BrCas12b enzyme, was subtracted from the sample fluorescence reading to generate background-subtracted fluorescence results. Linear regression analysis and enzyme dynamic analysis were conducted using GraphPad Prism 8.

2.8. Defining the minimum target range

We designed a series of truncated spacer 1 sequences (Supplementary Table S3) from the 3′ end and integrated them into the pMD19-T plasmid. The template was amplified from this recombinant plasmid by PCR using linearized primers. The PCR products were purified using the StarPrep PCR & DNA Fragment Purification Kit. The reaction mixture contained 8 nM Cas12b protein, 20 nM crRNA–tracrRNA complex, and 1 μg template DNA in cleavage buffer and was incubated at 62 °C for 1 h. The product was analyzed by electrophoresis on a 1.5% agarose gel containing SYBR Safe DNA Gel Stain.

2.9. Seed sequence identification

We used the spacer 1 sequence from the BrCas12b locus as the target sequence for design of a series of mutated sequences (Supplementary Tables S4, S5). These mutated sequences were integrated into the pMD19-T plasmid, which was then linearized and PCR amplified to prepare the templates. The PCR products ware purified using the StarPrep PCR & DNA Fragment Purification Kit. The reaction mixture contained 8 nM Cas12b protein, 20 nM crRNA–tracrRNA complex, and 1 μg template DNA in cleavage buffer and was incubated at 62 °C for 1 h. The product was analyzed by electrophoresis on a 1.5% agarose gel containing SYBR Safe DNA Gel Stain.

3. Results

3.1. Identification and structural analysis of BrCas12b

Brevibacillus sp. SYSU G02855 genome was found to contain a novel member of the Cas12b family, named BrCas12b, following traditional naming protocols. The CRISPR-BrCas12b locus is composed of a full CRISPR array containing 16 unique spacers separated by 36-nt long repeats and a cluster consisting of the *BrCas12b*, *Cas4-1*, and *Cas2* genes (Fig. 1A). This cluster exhibited gene overlap between the 5′ end of *Cas4-1* and 3′ end of *BrCas12b* (Fig. 1A). Analyses of the amino acid sequences showed that the BrCas12b protein had 34% sequence identity with the previously reported AacCas12b protein from *A. acidoterrestris* (Fig. S1). Surprisingly, the BrCas12b protein is smaller (1089 amino acids) than the previously discovered Cas12b proteins (AacCas12b: 1129 amino acids (aa); AaCas12b: 1129 aa; AmCas12b: 1146 aa; AkCas12a: 1147 aa) [28,35]. An alignment analysis predicts that the C-terminal portion of BrCas12b should adopt a mixed alpha helix/beta strand mode (Fig. S2) and contain an RuvC-like domain (Figs. S2, S3). Although BrCas12b exhibits a lower identity value than other reported Cas12b proteins, the computational analysis of the BrCas12b loci suggests that the novel homolog satisfies the structural characterization of the CRISPR-Cas12b system.

3.2. In vitro characterization of BrCas12b DNA cleavage activity

For biochemical characterization of BrCas12b, we synthesized the BrCas12b protein from the *Escherichia coli* BL21 heterogeneous expression system, which was purified to obtain 5 mg BrCas12b protein for further experiments (Fig. S4). All known type V-B CRISPR systems require dual-RNA (tracrRNA and crRNA)-guided endonuclease systems to achieve virus removal [28,35,36]. Using RNA sequencing, we confirmed that the crRNA consists of a 27-nt direct repeat and 38-nt spacer (Fig. S5A), and that the tracrRNA was significantly expressed and processed to 99 nt (Fig. S5B).

Given that the protospacer-adjacent motif (PAM) region plays a crucial role in target recognition, and that the PAM of the Cas12b family is located in the 5′ direction of the target DNA, we further explored the requirement for PAM in BrCas12b interference *via* an *in vitro* plasmid interference assay (Fig. 1B). We generated a PAM screening library of pMD19-T plasmids containing an inserted protospacer 1 sequence and an additional 7-nt random sequence upstream of the protospacer. Using a BrCas12b cleavage assay, we distinguished between cleaved plasmids with and those without addition of Cas12b, demonstrating that both TTN and TNT patterns were efficiently cleaved (Fig. 1C). Referencing the conditions of *in vitro* cleavage assay of AacCas12b, we found that purified BrCas12b with a tracrRNA–crRNA component was able to cleave target DNA containing a spacer 1 sequence (Fig. 1D), indicating that BrCas12b is a tracrRNA–crRNA-guided endonuclease that degrades both plasmid and linear DNA. Next, we determined the metal ion dependence of BrCas12b and found that BrCas12b activity was dependent on divalent metal ions, including Mg^{2+}, Ca^{2+}, Ba^{2+}, Co^{2+}, and Mn^{2+} (Fig. S6).

We further identified the cleavage site of BrCas12b from Sanger sequencing data of the dsDNA ends after cleavage. The results imply that BrCas12b cleaves dsDNAs, producing staggered ends at the distal position of the PAM. Specifically, cleavage occurred after the 16th base on the non-targeted (+) strand and after the 23rd base on the targeted (−) strand (Fig. 1E). This cleavage pattern differs from that of Cas9, which produces blunt ends, but is similar to that of the Cpf1 system.

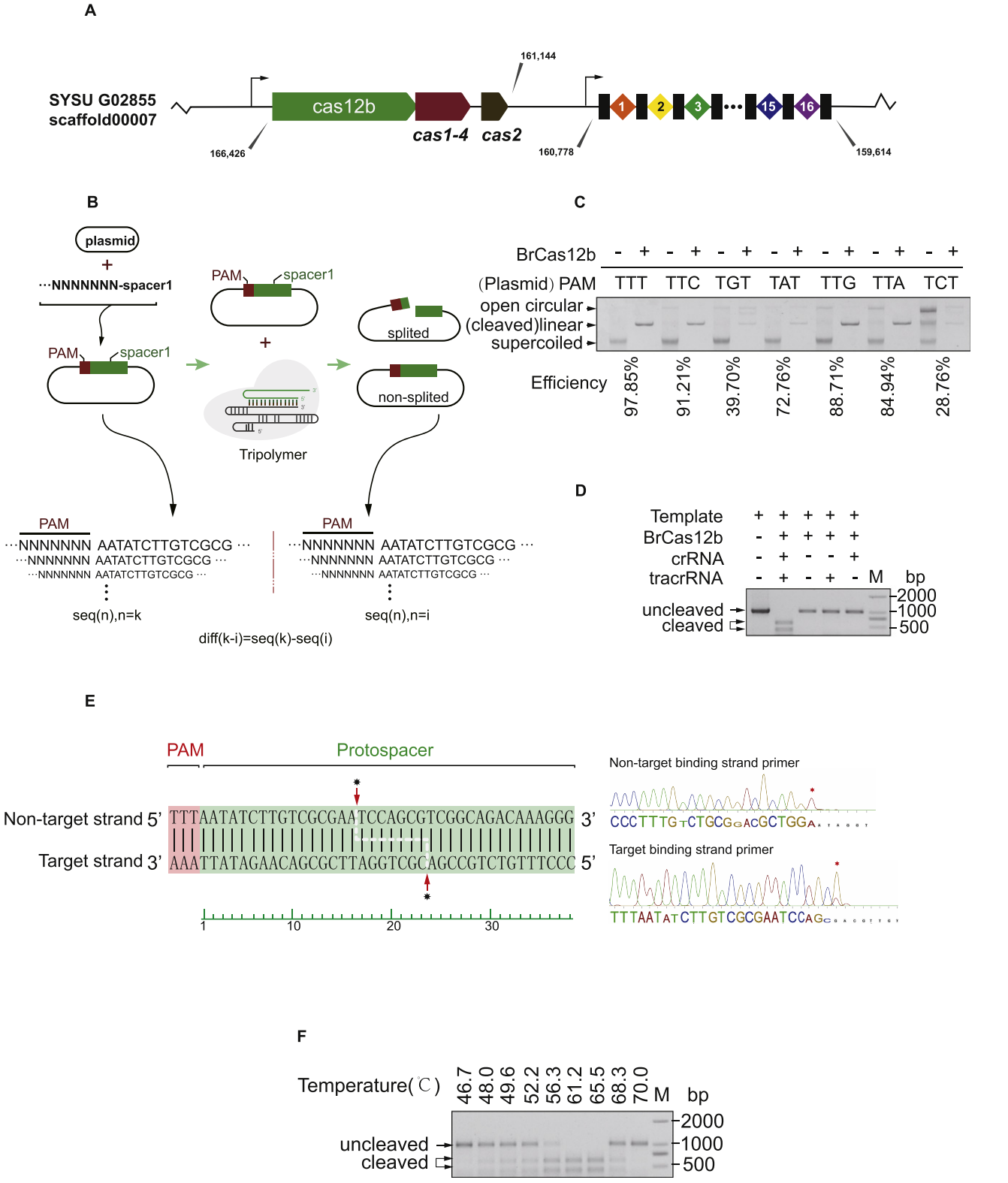


Fig. 1. BrCas12b cleaves dsDNA with dual-RNA *in vitro*. A, Schematic illustration of the genomic architecture of the CRISPR-Cas12b system from *Brevibacillus* sp. SYSU G02855. CRISPR-BrCas12b loci contains *cas12b* gene in blue, *cas 1-4* fusion in dark red, *cas2* in brown, spacers in various color separated by repeats in black. Genomic architectural drawing is scaled according to the gene size. B, Schematic of PAM identification screen. C, *In vitro* validation of the PAM requirements of BrCas12b indicating that PAM matching the 5′-TTT sequence can be robustly cleaved. D, *In vitro* cleavage of target with purified BrCas12b protein illustrates that the catalytic activity of BrCas12b is depends on dual-RNA. BrCas12b is incapable of cleaving target DNA without one of the both crRNA and tracrRNA. E, Identification of cleavage site both in target strand and non-target strand. Schematic represent a partial target DNA sequence and the Sanger sequencing results. The PAM sequence is shown in pink; protospacer sequence is represented in green. Cleavage sites mapped in Sanger sequencing data and marked with red arrows. Termination of primer extension in the sequencing reaction indicates the position of the cleavage site. F, *In vitro* temperature adaptability analysis of BrCas12b targeting shows that the optimal temperature range is between 48 °C and 68 °C. Especially, enzyme shows robuster cleavage activity between the 61 °C and 65 °C.

3.3. Active temperature range of BrCas12b

Compared with the reference temperature (55 °C), which is the optimal reaction temperature of the AacCas12b protein, we investigated the active temperature range of BrCas12b. The BrCas12b enzyme appeared to be active at temperatures above the active range of AacCas12b, as the bacterium *Brevibacillus* sp. SYSU G02855 was cultivated at higher temperatures than those for *A. acidoterrestris*. To validate this hypothesis, an assay of the active temperature range was performed to define the optimal DNA cleavage temperature of BrCas12b. Between 30 °C and 80 °C, the temperature assay indicated that BrCas12b maintains nuclease activity at higher temperatures compared with

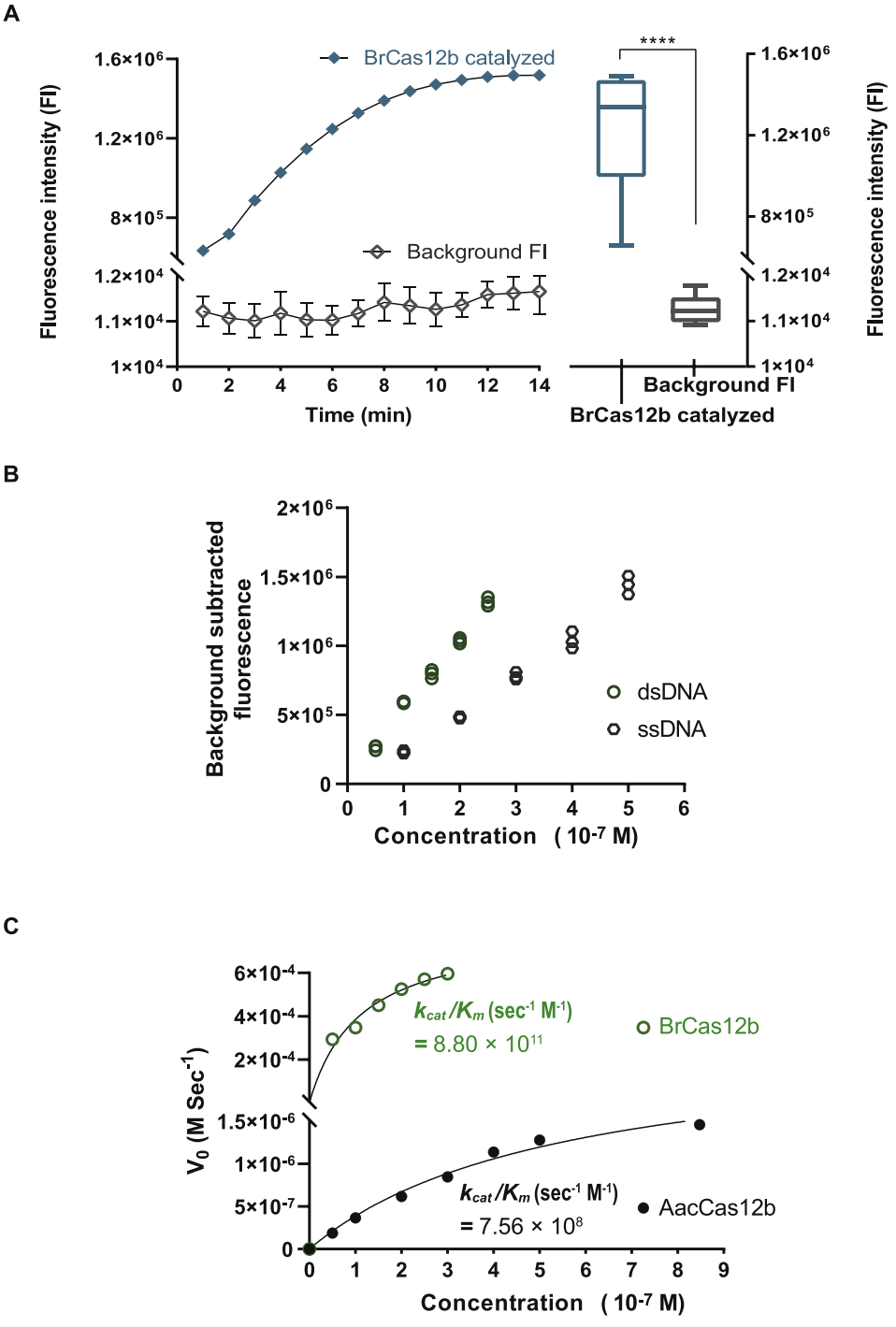


Fig. 2. Detection of Cas12b enzyme cleavage efficiency using a fluorescence resonance energy transfer-based assay. A, Evaluation of the validity of the fluorescence detection method. Blue diamond, fluorescence intensity representing BrCas12b catalysis; gray open diamond, fluorescence intensity without BrCas12b catalysis and background fluorescence intensity (FI); n = 3 technical replicates, two-tailed Student's *t*-test; $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$, $^{****}P < 0.0001$; bars represent the mean ± SEM. B, Linear regression analysis of two types of beacons modified with 6-FAM in the 5′ site and BHQ1 in the 3′ site. The background-subtracted fluorescence is the dependent variable (Y), and the concentration of beacons is the independent variable (X), yielding the following linear equations: $Y_{dsDNA} = 5.1 \times 10^{12} X + 35{,}420$, $R^2 = 0.9923$ and $Y_{ssDNA} = 2.9 \times 10^{12} X - 83{,}451$, $R^2 = 0.9899$. C, Kinetic analysis of BrCas12b and AacCas12b. Green open circle, initial velocity of the BrCas12b enzymatic reaction; black circle, initial velocity of the AacCas12b enzymatic reaction; k_{cat}, number of molecules of substrate produced per catalysis site per second; K_m, Michaelis–Menten constant; k_{cat}/K_m, specificity constant of catalytic efficiency.

AacCas12b (Figs. 1F, S8B), with maximal cleavage activity between 48 °C and 68 °C. Between 61 °C and 65 °C, BrCas12b cleavage activity was considerably greater than that of AacCas12b (Fig. 1F).

3.4. Robust DNA cleavage activity of thermophilic BrCas12b

The template cleavage efficiency of AacCas12b was significantly lower than that of BrCas12b based on the *in vitro* active temperature assay. However, the optical densities of the gel electrophoretic bands obtained in this assay may not accurately reflect the efficiency of Cas12b enzymes for cleaving target DNA, which is influenced by the gel concentration, DNA stain, imaging equipment, and other factors. To resolve this problem, we introduced and optimized a fluorometric assay method for accurate measurement of Cas12b activity, based on a conventional method used to investigate the kinetics of the Cas9 effector complex assembly [34].

In the fluorometric assay, a mixture of 7 nM BrCas12b and 10 nM crRNA, pre-annealed with equimolar tracrRNA, was added to a solution containing 250 nM beacons (dsDNA) and incubated at 62 °C for 15 min. The resulting fluorescence intensity was significantly higher in the BrCas12b-catalyzed group than the control group, which lacked the BrCas12b enzyme (Fig. 2A). We further evaluated the fluorescence intensity using two types of beacons, ssDNA and dsDNA, and found a linear relationship between the fluorescence intensity gradient and the concentration of beacons (Fig. 2B).

Using this fluorophore quencher reporter assay, we investigated the kinetics of the cleavage reaction using various substrates. We assembled BrCas12b with cr/tracrRNA and crRNA-complementary ssDNA (target strain), dsDNA, or non-target strain ssDNA. We used the AacCas12b system as a comparison. Both the dsDNA and target ssDNA groups exhibited robust increases in fluorescence intensity during the reaction (Fig. 2B). Analysis of the Michaelis–Menten reaction kinetics indicated that BrCas12b has a strong affinity for dsDNA. When the substrate was dsDNA, the k_{cat}/K_m value (8.80×10^{11} s^{-1} M^{-1}) was approximately three orders of magnitude greater for the BrCas12b system than the AacCas12b system (7.56×10^{8} s^{-1} M^{-1}) (Fig. 2C). We conclude that the novel BrCas12b system has robust target DNA cleavage efficiency, rendering it a promising tool for gene manipulation.

3.5. Defining the minimum target range

In our previous research, we adopted a conservative strategy in our target cleavage assays by using the full length of the spacer as the target region. Here, we investigated the minimum required target range by using truncated protospacers derived from a template (Fig. 3). We created a series of templates that were truncated in the 3′-to-5′ direction. The templates for various truncated protospacers were used in *in vitro* cleavage assays and the products were purified by electrophoresis. The cleavage efficiency decreased significantly when the protospacer was shortened to 16 nt. This contrasts with similar experiments with other Cas12b family members, specifically Cas12b from *A. acidoterrestris* and *Bacillus thermoamylovorans* (evolutionarily close to BrCas12b), which require a minimum of 20 nt of protospacer length.

3.6. Identification of the seed sequence from CRISPR-BrCas12b system

In the CRISPR-Cas system, DNA cleavage is dependent on recognition of the PAM-proximal sequence, referred to as the “seed sequence”, and introduction of a single point mutation in this sequence will abolish CRISPR-mediated immunity [37]. To determine whether target heterogeneity impacts BrCas12b cleavage activity, comparative analyses of *in vitro* interference by *A. acidoterrestris* Cas12b were performed. We

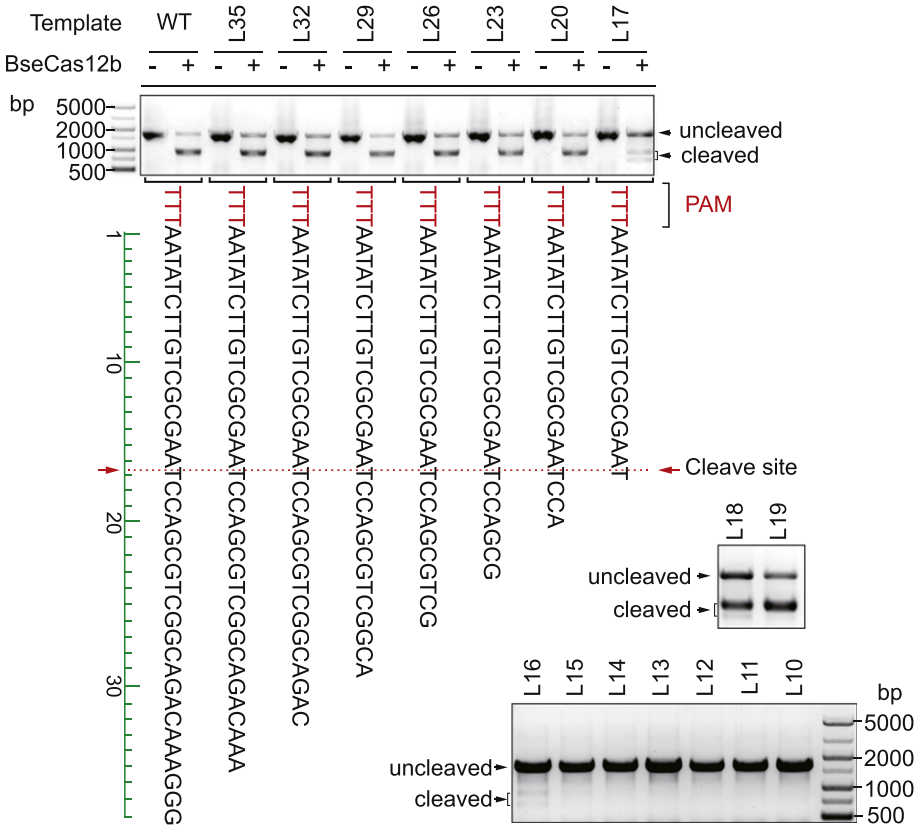


Fig. 3. Determination of the minimum target region. The targets contain multiple truncated protospacers which were truncated in the 3′-to-5′ direction. The cleavage products were loaded in Agarose gel and separated by electrophoresis.

designed and synthesized a series of single point mutations in the protospacer, which served as the substrate for cleavage activity assays of Cas12b from *A. acidoterrestris* and *Brevibacillus* sp. SYSU G02855. The results showed that BrCas12b required a small, rigorously consistent region of target DNA, with a length of only 4 nt, at a distance 3–6 nt downstream of the PAM (Fig. 4A). Surprisingly, both A1T and A2T substitutions in the protospacer, introducing a mismatch in the seed, decreased the efficiency of BrCas12b cleavage activity (Fig. 4B).

On this basis, we next investigated how the base composition at positions 1 and 2 in the protospacer affected the efficiency of target DNA cleavage. Substitutions A1C, A1G, A2C, and A2G significantly decreased the cleavage efficiency, particularly when the pyrimidine nucleotide was cytidine, but still allowed cleavage (Fig. 4B, C). By contrast, the AacCas12b complex strictly requires a 6-nt seed region immediately adjacent to the PAM site and is low-fault-tolerant (Fig. S8C). Substitution at the PAM-distal region position 16 (A16C and A16G) also significantly

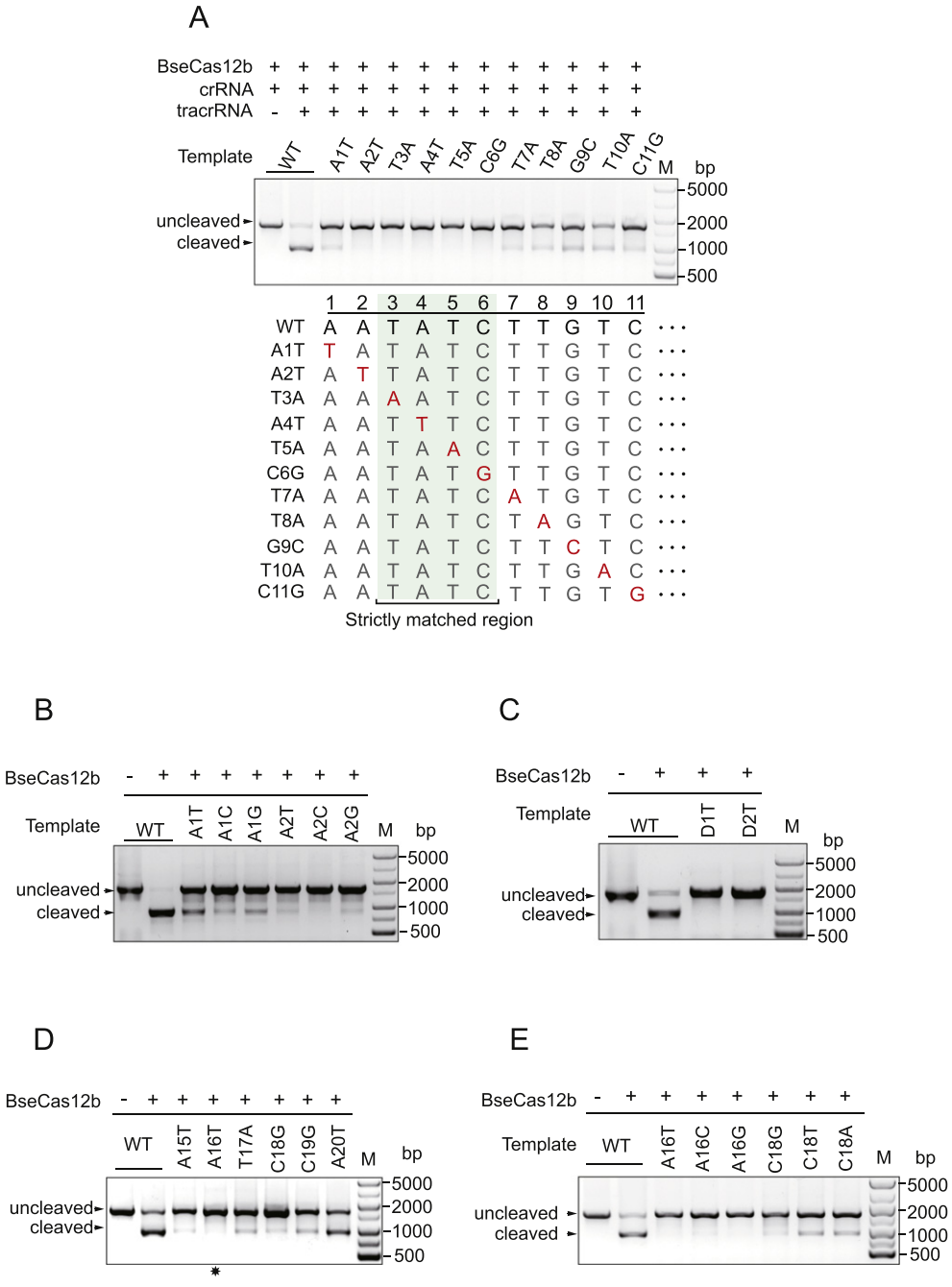


Fig. 4. Identifying the seed sequence of BrCas12b. A, *In vitro* cleavage of mutational target with a series of single point mutations in protospacer. The nucleotide sequences represent the mutational region. Mutated nucleotides are highlighted in red. The cleavage products were resolved by Agarose gel electrophoresis and visualized by SYBR-Safe staining. B, Determining the effects of substitutions at PAM adjoining range *in vitro* cleavage assay. C, Agarose gel electrophoresis analysis shows that the deletions of position 1 or 2 in protospacer block the target DNA cleavage. D, Analysis cleavage efficiency of single point mutations shows that the mutations both A16T and C18G can reduce cutting efficiency significantly. E, *In vitro* cleavage assay of double stranded DNAs containing substitution at the PAM-distal region position 16 and 18.

reduced the cleavage efficiency (Fig. 4D, E), presumably because of obstruction at the target DNA cleavage site.

4. Discussion

In this study, we identified a novel member of the V-B type CRISPR-Cas family from *Brevibacillus* sp. SYSU G02855 (BrCas12b), representing a new enzyme tool for genome editing and clinical diagnosis with interesting characteristics.

BrCas12b protein is a relatively compact protein (1089 amino acids), compared with SpCas9 and AsCas12a, suggesting that it may be delivered more easily into cells. Moreover, BrCas12b is derived from a non-pathogenic microorganism, unlike SpCas9 (from *Streptococcus pyogenes*) and SaCas9 (from *Staphylococcus aureus*) [38], offering a distinct safety advantage for therapeutic genome editing in human genetic diseases.

The BrCas12b enzyme is capable of binding and cleaving target DNA at high temperatures due to its stability, a feature that makes it suitable for development as a diagnostic tool. Various CRISPR effectors, including Cas13a, Cas12a and Csm6, have been developed into a series of clinical diagnostic platforms that depend on isothermal amplification to achieve initial amplification of specific target nucleic acids [39]. However, these diagnostic methods are executed in two steps, because the temperature requirement for template melting and that for mesophilic effector activity do not overlap. Therefore, a thermophilic enzyme active at template melting temperatures, such as BrCas12b, will simplify the operation of environmental and laboratory equipment and will be conducive to deployment in areas without laboratory infrastructure. This protein can realistically meet the demands of an ideal diagnostic test for use in the developing world [40].

Moreover, BrCas12b has a short seed sequence requirement, which we experimentally defined as a 4-nt sequence. This short sequence differs distinctly from the reported seed sequence of the AacCas12b enzyme (18 nt), although the two share 34% sequence identity. This result distinguishes BrCas12b from other types of CRISPR-Cas systems; *e.g.*, subtype I-B from *Haloferax volcanii* requires a 10-nt seed sequence [41], subtype I-E from *Escherichia coli* requires a 7-nt seed sequence [37], subtype I-F from *Pseudomonas aeruginosa* requires an 8-nt sequence [42], subtype II-A from *Streptococcus pneumonia* requires a 12-nt sequence [43], and subtype V-A from *Francisella novicida* requires a 5-nt sequence [25]. These differences indicating that the pattern recognition of BrCas12b targets is dissimilar to that of other effectors. Moreover, substitution at position 16 in the protospacer can rescue the target cleavage; perhaps mutation of the cleavage site destroys the R-loop complex. Such conservation is observed with other Cas12b family members, including AacCas12b and BthCas12b [44]. Even though the natural function of this difference remains unclear, we propose that thermophilic BrCas12b contains a specific pattern for target matching. Elucidation of the mechanism of this matching process might strongly increase cleavage specificity and reduce off-target activity in genome editing.

Although this work is based on systematic characterization of the novel BrCas12b protein, revealing its promising applications, one issue requires further investigation. As noted above, BrCas12b has strong DNase activity ($k_{cat}/K_m = 8.80 \times 10^{11}\ s^{-1}\ M^{-1}$) at its optimal temperature range, up to 62 °C, but this temperature is not suitable for mammalian genome editing. A similar dilemma was encountered with Cas12b protein from *Bacillus hisashii* (BhCas12b), and researchers utilized rational engineering of BhCas12b to address this limitation and increase its dsDNA cleavage activity at 37 °C [28]. Following this strategy, further experiments should be designed to promote adaption of the temperature range of BrCas12b for genome editing of mesophilic organisms. Moreover, in the process of rational engineering of BrCas12b, we can explore its thermophilic evolutionary basis and molecular mechanism.

In conclusion, our findings expand the epistemic range of the Cas12b family and fill the void of a thermophilic Cas12b enzyme. We believe that BrCas12b is an advantageous model for in-depth analysis of the high-temperature adaptability of the Cas12b family and its mechanism of targeting matches, and we anticipate that continued exploitation of the BrCas12b enzyme will provide a useful tool for genome editing and clinical diagnosis.

CRediT authorship contribution statement

Ye Tian: Conceptualization, Writing - original draft, Writing - review & editing. **Rui-Rui Liu:** Investigation, Writing - review & editing. **Wen-Dong Xian:** Investigation, Writing - review & editing. **Mengjie Xiong:** Investigation, Writing - review & editing. **Min Xiao:** Investigation, Writing - review & editing. **Wen-Jun Li:** Conceptualization, Writing - original draft, Writing - review & editing.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Acknowledgements

This work was financially supported by National Natural Science Foundation of China (Grant no. 91951205) and Key Realm R&D Program of GuangDong Province (Grant no. 2018B020206001).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2020.01.079.

References

[1] L.A. Marraffini, E.J. Sontheimer, CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea, Nat. Rev. Genet. 11 (3) (2010) 181–190, https://doi.org/10.1038/nrg2749.
[2] H. Deveau, J.E. Garneau, S. Moineau, CRISPR/Cas system and its role in phage-bacteria interactions, Annu. Rev. Microbiol. 64 (2010) 475–493, https://doi.org/10.1146/annurev.micro.112408.134123.
[3] R. Barrangou, L.A. Marraffini, CRISPR-Cas systems: prokaryotes upgrade to adaptive immunity, Mol. Cell 54 (2) (2014) 234–244, https://doi.org/10.1016/j.molcel.2014.03.011.
[4] J.E. Garneau, M.E. Dupuis, M. Villion, D.A. Romero, R. Barrangou, P. Boyaval, C. Fremaux, P. Horvath, A.H. Magadan, S. Moineau, The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA, Nature 468 (7320) (2010) 67–71, https://doi.org/10.1038/nature09523.
[5] B. Wiedenheft, S.H. Sternberg, J.A. Doudna, RNA-guided genetic silencing systems in bacteria and archaea, Nature 482 (7385) (2012) 331–338, https://doi.org/10.1038/nature10886.
[6] R. Barrangou, CRISPR-Cas systems and RNA-guided interference, Wiley Interdiscip. Rev. RNA 4 (3) (2013) 267–278, https://doi.org/10.1002/wrna.1159.
[7] L.A. Marraffini, CRISPR-Cas immunity in prokaryotes, Nature 526 (7571) (2015) 55–61, https://doi.org/10.1038/nature15386.
[8] G. Amitai, R. Sorek, CRISPR-Cas adaptation: insights into the mechanism of action, Nat. Rev. Microbiol. 14 (2) (2016) 67–76, https://doi.org/10.1038/nrmicro.2015.14.
[9] E. Deltcheva, K. Chylinski, C.M. Sharma, K. Gonzales, Y. Chao, Z.A. Pirzada, M.R. Eckert, J. Vogel, E. Charpentier, CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III, Nature 471 (7340) (2011) 602–607, https://doi.org/10.1038/nature09886.
[10] J.S. Chen, E. Ma, L.B. Harrington, X. Tian, J.A. Doudna, CRISPR-Cas12a target binding unleashes single-stranded DNase activity, Science 360 (6387) (2018) 436–439, https://doi.org/10.1126/science.aar6245.
[11] S.Y. Li, Q.X. Cheng, J.K. Liu, X.Q. Nie, G.P. Zhao, J. Wang, CRISPR-Cas12a has both cis- and trans-cleavage activities on single-stranded DNA, Cell Res. 28 (4) (2018) 491–493, https://doi.org/10.1038/s41422-018-0022-x.
[12] O. Niewoehner, C. Garcia-Doval, J.T. Rostol, C. Berk, F. Schwede, L. Bigler, J. Hall, L.A. Marraffini, M. Jinek, Type III CRISPR-Cas systems produce cyclic oligoadenylate second messengers, Nature 548 (7669) (2017) 543–548, https://doi.org/10.1038/nature23467.
[13] C. Rouillon, J.S. Athukoralage, S. Graham, S. Gruschow, M.F. White, Control of cyclic oligoadenylate synthesis in a type III CRISPR system, Elife 7 (2018) https://doi.org/10.7554/eLife.36734.
[14] K.S. Makarova, Y.I. Wolf, O.S. Alkhnbashi, F. Costa, S.A. Shah, S.J. Saunders, R. Barrangou, S.J. Brouns, E. Charpentier, D.H. Haft, P. Horvath, S. Moineau, F.J. Mojica, R.M. Terns, M.P. Terns, M.F. White, A.F. Yakunin, R.A. Garrett, J. van der Oost, R. Backofen, E.V. Koonin, An updated evolutionary classification of CRISPR-Cas systems, Nat. Rev. Microbiol. 13 (11) (2015) 722–736, https://doi.org/10.1038/nrmicro3569.

[15] D. Katrekar, M. Hu, P. Mali, Advances in CRISPR-Cas based genome engineering, Curr. Opin. Biomed. Eng. 1 (2017) 78–86, https://doi.org/10.1016/j.cobme.2017.04.001.
[16] K.M. Lewis, A. Ke, Building the class 2 CRISPR-Cas arsenal, Mol. Cell 65 (3) (2017) 377–379, https://doi.org/10.1016/j.molcel.2017.01.024.
[17] P. Mali, K.M. Esvelt, G.M. Church, Cas9 as a versatile tool for engineering biology, Nat. Methods 10 (10) (2013) 957–963, https://doi.org/10.1038/nmeth.2649.
[18] P.D. Hsu, E.S. Lander, F. Zhang, Development and applications of CRISPR-Cas9 for genome engineering, Cell 157 (6) (2014) 1262–1278, https://doi.org/10.1016/j.cell.2014.05.010.
[19] T. Jakociunas, M.K. Jensen, J.D. Keasling, CRISPR/Cas9 advances engineering of microbial cell factories, Metab. Eng. 34 (2016) 44–59, https://doi.org/10.1016/j.ymben.2015.12.003.
[20] K. Zhang, R. Deng, Y. Li, L. Zhang, J. Li, Cas9 cleavage assay for pre-screening of sgRNAs using nicking triggered isothermal amplification, Chem. Sci. 7 (8) (2016) 4951–4957, https://doi.org/10.1039/c6sc01355d.
[21] W. Zhou, L. Hu, L. Ying, Z. Zhao, P.K. Chu, X.F. Yu, A CRISPR-Cas9-triggered strand displacement amplification method for ultrasensitive DNA detection, Nat. Commun. 9 (1) (2018) 5012, https://doi.org/10.1038/s41467-018-07324-5.
[22] H. Kim, S.T. Kim, J. Ryu, B.C. Kang, J.S. Kim, S.G. Kim, CRISPR/Cpf1-mediated DNA-free plant genome editing, Nat. Commun. 8 (2017), 14406. https://doi.org/10.1038/ncomms14406.
[23] X. Tang, L.G. Lowder, T. Zhang, A.A. Malzahn, X. Zheng, D.F. Voytas, Z. Zhong, Y. Chen, Q. Ren, Q. Li, E.R. Kirkland, Y. Zhang, Y. Qi, A CRISPR-Cpf1 system for efficient genome editing and transcriptional repression in plants, Nat. Plants 3 (2017), 17018. https://doi.org/10.1038/nplants.2017.18.
[24] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J.A. Doudna, E. Charpentier, A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity, Science 337 (6096) (2012) 816–821, https://doi.org/10.1126/science.1225829.
[25] B. Zetsche, J.S. Gootenberg, O.O. Abudayyeh, I.M. Slaymaker, K.S. Makarova, P. Essletzbichler, S.E. Volz, J. Joung, J. van der Oost, A. Regev, E.V. Koonin, F. Zhang, Cpf1 is a single RNA-guided endonuclease of a class 2 CRISPR-Cas system, Cell 163 (3) (2015) 759–771, https://doi.org/10.1016/j.cell.2015.09.038.
[26] S. Shmakov, O.O. Abudayyeh, K.S. Makarova, Y.I. Wolf, J.S. Gootenberg, E. Semenova, L. Minakhin, J. Joung, S. Konermann, K. Severinov, F. Zhang, E.V. Koonin, Discovery and functional characterization of diverse class 2 CRISPR-Cas systems, Mol. Cell 60 (3) (2015) 385–397, https://doi.org/10.1016/j.molcel.2015.10.008.
[27] L. Liu, P. Chen, M. Wang, X. Li, J. Wang, M. Yin, Y. Wang, C2c1-sgRNA complex structure reveals RNA-guided DNA cleavage mechanism, Mol. Cell 65 (2) (2017) 310–322, https://doi.org/10.1016/j.molcel.2016.11.040.
[28] J. Strecker, S. Jones, B. Koopal, J. Schmid-Burgk, B. Zetsche, L. Gao, K.S. Makarova, E.V. Koonin, F. Zhang, Engineering of CRISPR-Cas12b for human genome editing, Nat. Commun. 10 (1) (2019) 212, https://doi.org/10.1038/s41467-018-08224-4.
[29] J. Besemer, A. Lomsadze, M. Borodovsky, GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions, Nucleic Acids Res. 29 (12) (2001) 2607–2618, https://doi.org/10.1093/nar/29.12.2607.
[30] R.C. Edgar, MUSCLE: multiple sequence alignment with high accuracy and high throughput, Nucleic Acids Res. 32 (5) (2004) 1792–1797, https://doi.org/10.1093/nar/gkh340.
[31] M.N. Price, P.S. Dehal, A.P. Arkin, FastTree 2 – approximately maximum-likelihood trees for large alignments, PLoS One 5 (3) (2010), e9490. https://doi.org/10.1371/journal.pone.0009490.
[32] A. Drozdetskiy, C. Cole, J. Procter, G.J. Barton, JPred4: a protein secondary structure prediction server, Nucleic Acids Res. 43 (W1) (2015) W389–W394, https://doi.org/10.1093/nar/gkv332.
[33] H. Li, R. Durbin, Fast and accurate short read alignment with Burrows–Wheeler transform, Bioinformatics 25 (14) (2009) 1754–1760, https://doi.org/10.1093/bioinformatics/btp324.
[34] V. Mekler, L. Minakhin, K. Severinov, Mechanism of duplex DNA destabilization by RNA-guided Cas9 nuclease during target interrogation, Proc. Natl. Acad. Sci. U. S. A. 114 (21) (2017) 5443–5448, https://doi.org/10.1073/pnas.1619926114.
[35] F. Teng, T. Cui, G. Feng, L. Guo, K. Xu, Q. Gao, T. Li, J. Li, Q. Zhou, W. Li, Repurposing CRISPR-Cas12b for mammalian genome engineering, Cell Discov. 4 (2018) 63, https://doi.org/10.1038/s41421-018-0069-3.
[36] E.V. Koonin, K.S. Makarova, F. Zhang, Diversity, classification and evolution of CRISPR-Cas systems, Curr. Opin. Microbiol. 37 (2017) 67–78, https://doi.org/10.1016/j.mib.2017.05.008.
[37] E. Semenova, M.M. Jore, K.A. Datsenko, A. Semenova, E.R. Westra, B. Wanner, J. van der Oost, S.J. Brouns, K. Severinov, Interference by clustered regularly interspaced short palindromic repeat (CRISPR) RNA is governed by a seed sequence, Proc. Natl. Acad. Sci. U. S. A. 108 (25) (2011) 10098–10103, https://doi.org/10.1073/pnas.1104144108.
[38] C.T. Charlesworth, P.S. Deshpande, D.P. Dever, J. Camarena, V.T. Lemgart, M.K. Cromer, C.A. Vakulskas, M.A. Collingwood, L. Zhang, N.M. Bode, M.A. Behlke, B. Dejene, B. Cieniewicz, R. Romano, B.J. Lesch, N. Gomez-Ospina, S. Mantri, M. Pavel-Dinu, K.I. Weinberg, M.H. Porteus, Identification of preexisting adaptive immunity to Cas9 proteins in humans, Nat. Med. 25 (2) (2019) 249–254, https://doi.org/10.1038/s41591-018-0326-x.
[39] J.S. Gootenberg, O.O. Abudayyeh, M.J. Kellner, J. Joung, J.J. Collins, F. Zhang, Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6, Science 360 (6387) (2018) 439–444, https://doi.org/10.1126/science.aaq0179.
[40] M. Urdea, L.A. Penny, S.S. Olmsted, M.Y. Giovanni, P. Kaspar, A. Shepherd, P. Wilson, C.A. Dahl, S. Buchsbaum, G. Moeller, D.C. Hay Burgess, Requirements for high impact diagnostics in the developing world, Nature 444 (Suppl. 1) (2006) 73–79, https://doi.org/10.1038/nature05448.
[41] L.K. Maier, S.J. Lange, B. Stoll, K.A. Haas, S. Fischer, E. Fischer, E. Duchardt-Ferner, J. Wohnert, R. Backofen, A. Marchfelder, Essential requirements for the detection and degradation of invaders by the Haloferax volcanii CRISPR/Cas system I-B, RNA Biol. 10 (5) (2013) 865–874, https://doi.org/10.4161/rna.24282.
[42] B. Wiedenheft, E. van Duijn, J.B. Bultema, S.P. Waghmare, K. Zhou, A. Barendregt, W. Westphal, A.J. Heck, E.J. Boekema, M.J. Dickman, J.A. Doudna, RNA-guided complex from a bacterial immune system enhances target recognition through seed sequence interactions, Proc. Natl. Acad. Sci. U. S. A. 108 (25) (2011) 10092–10097, https://doi.org/10.1073/pnas.1102716108.
[43] W. Jiang, D. Bikard, D. Cox, F. Zhang, L.A. Marraffini, RNA-guided editing of bacterial genomes using CRISPR-Cas systems, Nat. Biotechnol. 31 (3) (2013) 233–239, https://doi.org/10.1038/nbt.2508.
[44] I. Jain, L. Minakhin, V. Mekler, V. Sitnik, N. Rubanova, K. Severinov, E. Semenova, Defining the seed sequence of the Cas12b CRISPR-Cas effector complex, RNA Biol. 16 (4) (2019) 413–422, https://doi.org/10.1080/15476286.2018.1495492.