Electronic Supplementary Material (ESI) for RSC Advances.

Supplementary Information

An insight into the protospacer adjacent motif of *Streptococcus pyogenes* Cas9 with artificially stimulated RNA-guided-Cas9 DNA cleavage flexibility

Yujie Geng[a], Zixin Deng[a], Yuhui Sun*[a]

[a] Key Laboratory of Combinatorial Biosynthesis and Drug Discovery (Wuhan University), Ministry of Education, and Wuhan University School of Pharmaceutical Sciences, Wuhan 430071, People's Republic of China

* Correspondence and requests for materials should be addressed to E-mail: yhsun@whu.edu.cn

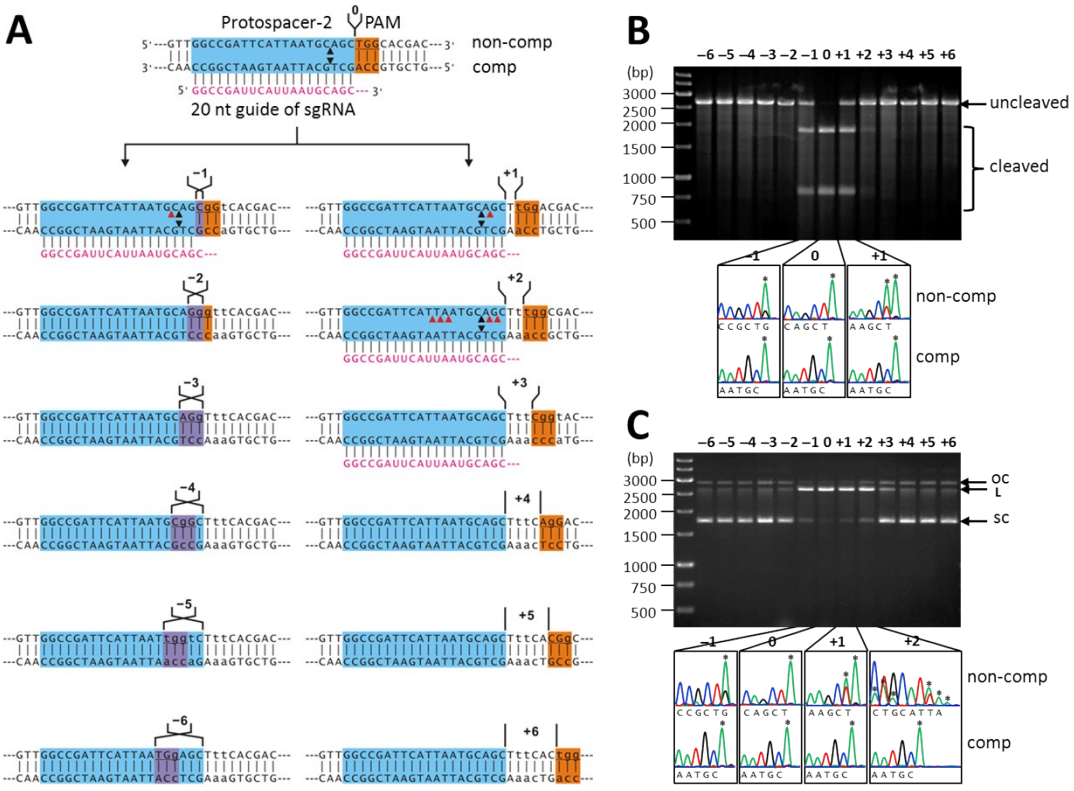

Fig. S1. Impact of length of the interspace between protospacer-2 and NGG on SpCas9 cleavage. (A) Schematic representation of dsDNA with interspace of variable length. Protospacer region, NGG (underlined) region and the overlapping region of NGG and protospacer are highlighted in blue, orange and purple, respectively, and sgRNA guide sequence is coloured red. The black and red triangles indicate known SpCas9 cleavage sites and newly observed cleavage sites in this study, respectively. The number with/without + and – indicate the number of base pairs of interspace or overlapping sequence between the protospacer and NGG, respectively. (B) and (C) show the cleavage results associated with dsDNA shown in (A) and end-sequencing analysis of cleaved products of *Ssp*I-linearized plasmid and circular plasmid, respectively. The 3' terminal A overhang, which is an artifact of the sequencing reaction, is represented by an asterisk. The following

abbreviations were used: OC, open circular DNA; L, linear DNA; SC, supercoiled DNA; comp, complementary strand; non-comp, non-complementary strand.

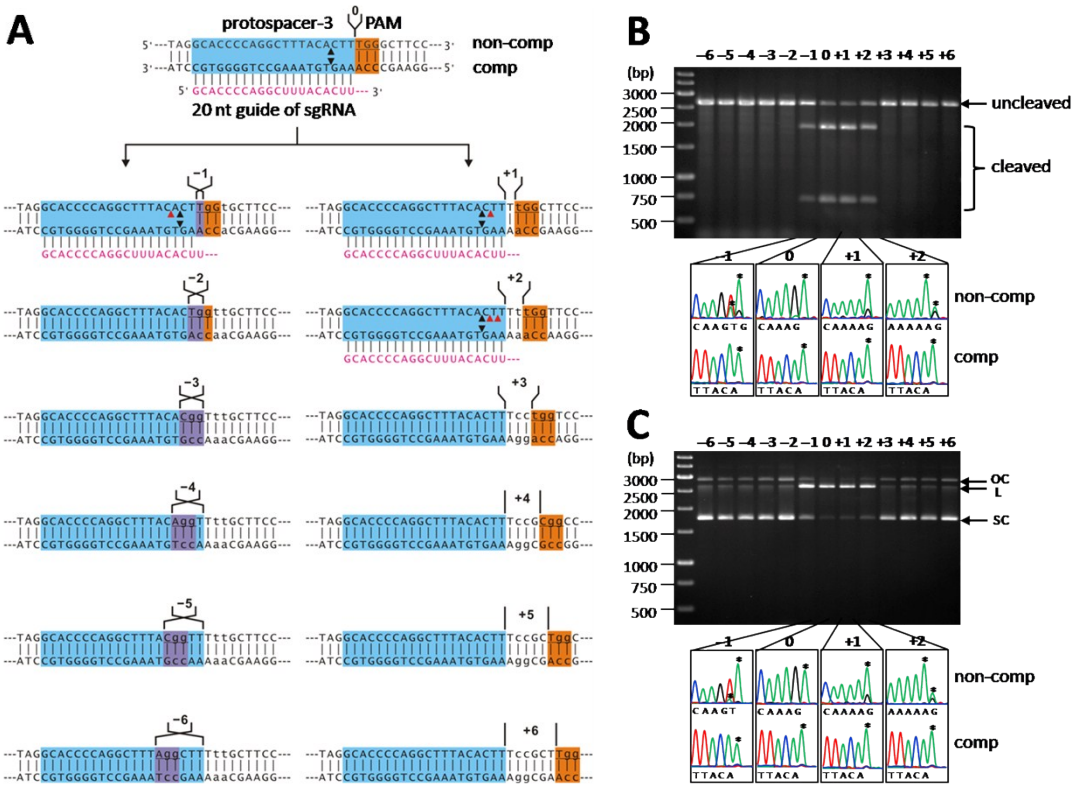

Fig. S2. Impact of length of the interspace between protospacer-3 and NGG on SpCas9 cleavage. (A) Schematic representation of dsDNA with interspace of variable length. The protospacer region, NGG (underlined) region and the overlapping region of NGG and protospacer are highlighted in blue, orange and purple, respectively, and sgRNA guide sequence is coloured red. The black and red indicate known cleavage sites of SpCas9 and cleavage sites that were observed for the first time as part of this study, respectively. The number with/without + and – indicate the number of base pairs associated with interspace or overlapping sequence between the

protospacer and NGG, respectively. (B) and (C) represent the cleavage results from the dsDNA shown in (A) and end-sequencing analysis of the cleavage products of *Ssp*I-linearized plasmid and circular plasmid, respectively. The 3' terminal A overhang which is an artifact of the sequencing reaction is represented by an asterisk. The following abbreviations were used: OC, open circular DNA; L, linear DNA; SC, supercoiled DNA; comp, complementary strand; non-comp, non-complementary strand.

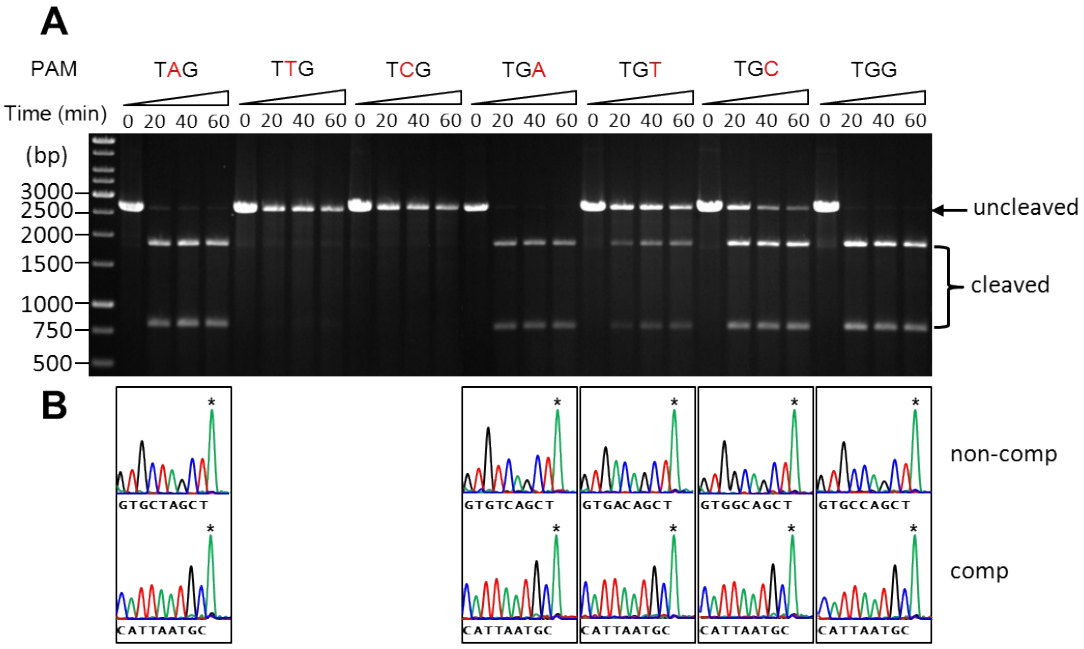

Fig. S3. The influence of single guanosine mutations in the NGG sequence of PAM on SpCas9-mediated cleavage using *Ssp*I-linearized plasmid on protospacer-2. Cleavage results (A) and end-sequencing results of cleavage products generated by treatment with SpCas9 for 60 min (B).

Table S1. Oligonucleotides used in this study

Oligonucleotides (5'→3') used for *in vitro* transcription templates		
protospacer-1	N1P1	GGAATTGTGAGCGGAGAAGAGTTTTAGAGCTAGAAATAGC
	N1P3	GATCACTAATACGACTCACTATAGGAATTGTGAGCGGAGAAGA
	N1-2P1	GGAATTGTGAGCGGAGAAAGGTTTTAGAGCTAGAAATAGC
	N1-2P3	GATCACTAATACGACTCACTATAGGAATTGTGAGCGGAGAAAG
	N1-3P1	GGAATTGTGAGCGGAGAAGGGTTTTAGAGCTAGAAATAGC
	N1-3P3	GATCACTAATACGACTCACTATAGGAATTGTGAGCGGAGAAGG
protospacer-2	N2P1	GGCCGATTCATTAATGCAGCGTTTTAGAGCTAGAAATAGC
	N2P3	GATCACTAATACGACTCACTATAGGCCGATTCATTAATGCAGC
	N2-2P1	GGCCGATTCATTAATGCAGGGTTTTAGAGCTAGAAATAGC
	N2-2P3	GATCACTAATACGACTCACTATAGGCCGATTCATTAATGCAGG
	N2-4P1	GGCCGATTCATTAATGCGGCGTTTTAGAGCTAGAAATAGC
	N2-4P3	GATCACTAATACGACTCACTATAGGCCGATTCATTAATGCGGC
	N2-5P1	GGCCGATTCATTAATTGGTCGTTTTAGAGCTAGAAATAGC
	N2-5P3	GATCACTAATACGACTCACTATAGGCCGATTCATTAATTGGTC
	N2-6P1	GGCCGATTCATTAATGGAGCGTTTTAGAGCTAGAAATAGC
	N2-6P3	GATCACTAATACGACTCACTATAGGCCGATTCATTAATGGAGC
protospacer-3	N3P1	GCACCCCAGGCTTTACACTTGTTTTAGAGCTAGAAATAGC
	N3P3	GATCACTAATACGACTCACTATAGCACCCCAGGCTTTACACTT
	N3-2P1	GCACCCCAGGCTTTACACTGGTTTTAGAGCTAGAAATAGC
	N3-2P3	GATCACTAATACGACTCACTATAGCACCCCAGGCTTTACACTG
	N3-3P1	GCACCCCAGGCTTTACACGGGTTTTAGAGCTAGAAATAGC
	N3-3P3	GATCACTAATACGACTCACTATAGCACCCCAGGCTTTACACGG
	N3-4P1	GCACCCCAGGCTTTACAGGTGTTTTAGAGCTAGAAATAGC
	N3-4P3	GATCACTAATACGACTCACTATAGCACCCCAGGCTTTACAGGT
	N3-5P1	GCACCCCAGGCTTTACGGTTGTTTTAGAGCTAGAAATAGC
	N3-5P3	GATCACTAATACGACTCACTATAGCACCCCAGGCTTTACGGTT
	N3-6P1	GCACCCCAGGCTTTAGGCTTGTTTTAGAGCTAGAAATAGC
	N3-6P3	GATCACTAATACGACTCACTATAGCACCCCAGGCTTTAGGCTT
	P2-4	AGCACCGACTCGGTGCCACTTTTCCAAGTTGATAACGGACTAGCCTTATTTTAACT
Oligonucleotides (5'→3') used for site-directed mutagenesis		
protospacer-1	N1TGG-S	TGTGAGCGGAGAAGATGGTCACACAGGAAACAGCTATGAC
	N1TGG-A	TCCTGTGTGACCATCTTCTCCGCTCACAATTCCACACAAC
	N1+1-S	TGTGAGCGGAGAAGATTGGCACACAGGAAACAGCTATGAC
	N1+1-A	TTCCTGTGTGCCAATCTTCTCCGCTCACAATTCCACACAA

	N1+2-S	TGTGAGCGGAGAAGATTTGGACACAGGAAACAGCTATGAC
	N1+2-A	TTTCCTGTGTCCAAATCTTCTCCGCTCACAATTCCACACA
	N1+3-S	TGTGAGCGGAGAAGATTTTGGCACAGGAAACAGCTATGAC
	N1+3-A	GTTTCCTGTGCCAAAATCTTCTCCGCTCACAATTCCACAC
	N1+4-S	TGTGAGCGGAGAAGAATTTTGGACAGGAAACAGCTATGAC
	N1+4-A	TGTTTCCTGTCCAAAATTCTTCTCCGCTCACAATTCCACA
protospacer-1	N1+5-S	TGTGAGCGGAGAAGAATTTTTGGCAGGAAACAGCTATGAC
	N1+5-A	CTGTTTCCTGCCAAAAATTCTTCTCCGCTCACAATTCCAC
	N1-1-S	TGTGAGCGGAGAAGAGGTTCACACAGGAAACAGCTATGAC
	N1-1-A	CCTGTGTGAACCTCTTCTCCGCTCACAATTCCACACAACA
	N1-2-S	TGTGAGCGGAGAAAGGTTTCACACAGGAAACAGCTATGAC
	N1-2-A	CTGTGTGAAACCTTTCTCCGCTCACAATTCCACACAACAT
	N1-3-S	TGTGAGCGGAGAAGGTTTTCACACAGGAAACAGCTATGAC
	N1-3-A	CTGTGTGAAAACCTTCTCCGCTCACAATTCCACACAACAT
protospacer-2	N2+1-S	TTAATGCAGCTTGGACGACAGGTTTCCCGACTGGAAAGCG
	N2+1-A	AACCTGTCGTCCAAGCTGCATTAATGAATCGGCCAACGCG
	N2+2-S	TAATGCAGCTTTGGCGACAGGTTTCCCGACTGGAAAGCGG
	N2+2-A	AAACCTGTCGCCAAAGCTGCATTAATGAATCGGCCAACGC
	N2+3-S	TAATGCAGCTTTCGGTACAGGTTTCCCGACTGGAAAGCGG
	N2+3-A	GGAAACCTGTACCGAAAGCTGCATTAATGAATCGGCCAAC
	N2+4-S	TAATGCAGCTTTCAGGACAGGTTTCCCGACTGGAAAGCGG
	N2+4-A	GAAACCTGTCCTGAAAGCTGCATTAATGAATCGGCCAACG
	N2+5-S	TAATGCAGCTTTCACGGCAGGTTTCCCGACTGGAAAGCGG
	N2+5-A	GGGAAACCTGCCGTGAAAGCTGCATTAATGAATCGGCCAA
	N2+6-S	TAATGCAGCTTTCACTGGAGGTTTCCCGACTGGAAAGCGG
	N2+6-A	CGGGAAACCTCCAGTGAAAGCTGCATTAATGAATCGGCCA
	N2-1-S	ATTAATGCAGCGGTCACGACAGGTT
	N2-1-A	AACCTGTCGTGACCGCTGCATTAAT
	N2-2-S	CGATTCATTAATGCAGGGTTCACGACAGGTTTCCC
	N2-2-A	GGGAAACCTGTCGTGAACCCTGCATTAATGAATCG
	N2-3-S	ATTAATGCAGGTTTCACGACAGGTTTCCCGACTGGAAAGC
	N2-3-A	ACCTGTCGTGAAACCTGCATTAATGAATCGGCCAACGCGC
	N2-4-S	TCATTAATGCGGCTTTCACGACAGGTTTCCCGACTGGAAA
	N2-4-A	ACCTGTCGTGAAAGCCGCATTAATGAATCGGCCAACGCGC
	N2-5-S	ATTCATTAATTGGTCTTTCACGACAGGTTTCCCGACTGGA
	N2-5-A	ACCTGTCGTGAAAGACCAATTAATGAATCGGCCAACGCGC
	N2-6-S	TTCATTAATGGAGCTTTCACGACAGGTTTCCCGACTGGAA
	N2-6-A	ACCTGTCGTGAAAGCTCCATTAATGAATCGGCCAACGCGC
	N2TAG-S	ATTAATGCAGCTAGCACGACAGGTT
	N2TAG-A	AACCTGTCGTGCTAGCTGCATTAAT
	N2TTG-S	ATTAATGCAGCTTGCACGACAGGTT
	N2TTG-A	AACCTGTCGTGCAAGCTGCATTAAT
	N2TCG-S	TAATGCAGCTCGCACGACAGGTTTCCCGACTGGAAAGCGG

	N2TCG-A	CCTGTCGTGCGAGCTGCATTAATGAATCGGCCAACGCGCG
	N2TGA-S	ATTAATGCAGCTGACACGACAGGTT
	N2TGA-A	AACCTGTCGTGTCAGCTGCATTAAT
	N2TGT-S	ATTAATGCAGCTGTCACGACAGGTT
	N2TGT-A	AACCTGTCGTGACAGCTGCATTAAT
	N2TGC-S	ATTAATGCAGCTGCCACGACAGGTT
protospacer-2	N2TGC-A	AACCTGTCGTGGCAGCTGCATTAAT
	N2TAA-S	TAATGCAGCTAACACGACAGGTTTCCCGACTGGAAAGCGG
	N2TAA-A	ACCTGTCGTGTTAGCTGCATTAATGAATCGGCCAACGCGC
	N2TAT-S	TAATGCAGCTATCACGACAGGTTTCCCGACTGGAAAGCGG
	N2TAT-A	ACCTGTCGTGATAGCTGCATTAATGAATCGGCCAACGCGC
	N2TAC-S	TAATGCAGCTACCACGACAGGTTTCCCGACTGGAAAGCGG
	N2TAC-A	ACCTGTCGTGGTAGCTGCATTAATGAATCGGCCAACGCGC
	N2TTA-S	TAATGCAGCTTACACGACAGGTTTCCCGACTGGAAAGCGG
	N2TTA-A	ACCTGTCGTGTAAGCTGCATTAATGAATCGGCCAACGCGC
	N2TTT-S	ATTAATGCAGCTTTCACGACAGGTT
	N2TTT-A	AACCTGTCGTGAAAGCTGCATTAAT
	N2TTC-S	TAATGCAGCTTCCACGACAGGTTTCCCGACTGGAAAGCGG
	N2TTC-A	ACCTGTCGTGGAAGCTGCATTAATGAATCGGCCAACGCGC
	N2TCA-S	TAATGCAGCTCACACGACAGGTTTCCCGACTGGAAAGCGG
	N2TCA-A	ACCTGTCGTGTGAGCTGCATTAATGAATCGGCCAACGCGC
	N2TCT-S	TAATGCAGCTCTCACGACAGGTTTCCCGACTGGAAAGCGG
	N2TCT-A	ACCTGTCGTGAGAGCTGCATTAATGAATCGGCCAACGCGC
	N2TCC-S	ATTAATGCAGCTCCCACGACAGGTT
	N2TCC-A	AACCTGTCGTGGGAGCTGCATTAAT
protospacer-3	N3TGG-S	TTTACACTTTGGTCTTCCGGCTCGTATGTTGTGTGGAATT
	N3TGG-A	GAGCCGGAAGACCAAAGTGTAAAGCCTGGGGTGCCTAATG
	N3+1-S	TTTACACTTTTGGCTTCCGGCTCGTATGTTGTGTGGAATT
	N3+1-A	AGCCGGAAGCCAAAAGTGTAAAGCCTGGGGTGCCTAATGA
	N3+2-S	TTTACACTTTTTGGTTCCGGCTCGTATGTTGTGTGGAATT
	N3+2-A	CGAGCCGGAACCAAAAAGTGTAAAGCCTGGGGTGCCTAAT
	N3+3-S	TTTACACTTTCCTGGTCCGGCTCGTATGTTGTGTGGAATT
	N3+3-A	ACGAGCCGGACCAGGAAAGTGTAAAGCCTGGGGTGCCTAA
	N3+4-S	TTTACACTTTCCGCGGCCGGCTCGTATGTTGTGTGGAATT
	N3+4-A	TACGAGCCGGCCGCGGAAAGTGTAAAGCCTGGGGTGCCTA
	N3+5-S	TTTACACTTTCCGCTGGCGGCTCGTATGTTGTGTGGAATT
	N3+5-A	ATACGAGCCGCCAGCGGAAAGTGTAAAGCCTGGGGTGCCT
	N3+6-S	TTTACACTTTCCGCTTGGGGCTCGTATGTTGTGTGGAATT
	N3+6-A	CATACGAGCCCCAAGCGGAAAGTGTAAAGCCTGGGGTGCC
	N3-1-S	CTTTACACTTGGTGCTTCCGGCTCGTATGTTGTGTGGAAT
	N3-1-A	GCCGGAAGCACCAAGTGTAAAGCCTGGGGTGCCTAATGAG
	N3-2-S	GCTTTACACTGGTTGCTTCCGGCTCGTATGTTGTGTGGAA
	N3-2-A	GCCGGAAGCAACCAGTGTAAAGCCTGGGGTGCCTAATGAG

	N3-3-S	GGCTTTACACGGTTTGCTTCCGGCTCGTATGTTGTGTGGA
	N3-3-A	GCCGGAAGCAAACCGTGTAAAGCCTGGGGTGCCTAATGAG
	N3-4-S	AGGCTTTACAGGTTTTGCTTCCGGCTCGTATGTTGTGTGG
	N3-4-A	GCCGGAAGCAAAACCTGTAAAGCCTGGGGTGCCTAATGAG
	N3-5-S	CAGGCTTTACGGTTTTTGCTTCCGGCTCGTATGTTGTGTG
	N3-5-A	GCCGGAAGCAAAAACCGTAAAGCCTGGGGTGCCTAATGAG
protospacer-3	N3-6-S	CCAGGCTTTAGGCTTTTTGCTTCCGGCTCGTATGTTGTGT
	N3-6-A	GCCGGAAGCAAAAAGCCTAAAGCCTGGGGTGCCTAATGAG
Oligonucleotides (5'→3') used as PAMmers at protospacer-2		
N2PL		TGGCACGACAGGTTTCCCG
N2PTTC		TTCCACGACAGGTTTCCCG
N2PM		AATGCAGCTGGCACGACAG
N2PR		GATTCATTAATGCAGCTGG
N2PE		AATGCAGCTGGCACGACAGGTTTCCCG
N2PA		CGGGAAACCTGTCGTGGGA