

SUPPLEMENTARY INFORMATION

Crystal structures of monomeric BsmI restriction endonuclease reveal coordinated sequential cleavage of two DNA strands

Rémi Sieskind¹, Sophia Missouri¹, Clément Madru¹, Isciane Commenge¹, Germain Niogret², Marcel Hollenstein², Yannick Rondelez³, Ludovic Sauguet¹, Ahmed Haouz⁴, Pierre Legrand⁵, Marc Delarue^{1*}

¹Institut Pasteur, Université Paris Cité, CNRS UMR 3528, Unit of Architecture and Dynamics of Biological Macromolecules, 75724 Paris, France

²Institut Pasteur, Université Paris Cité, CNRS UMR 3523, Unit of Bioorganic Chemistry of Nucleic Acids, 75724 Paris, France

³Laboratoire Gulliver, UMR7083 CNRS/ESPCI Paris-PSL Research University, 10 rue Vauquelin, 75005, Paris, France

⁴Institut Pasteur, Université Paris Cité, CNRS UMR 3528, Plateforme de cristallographie-C2RT, 75724 Paris, France

⁵Synchrotron SOLEIL, 91192 Gif-sur-Yvette, France

* To whom correspondence should be addressed. Tel: +33 1 45 68 86 05; Fax: +33 1 45 68 86 96; Email: delarue@pasteur.fr

Figure S1. 2 Analytical Ultra-Centrifugation results. The peak #1 corresponds to the monomeric apo-enzyme. The peak #2 corresponds to free DNA and the peak #3 to the Enzyme-DNA complex.

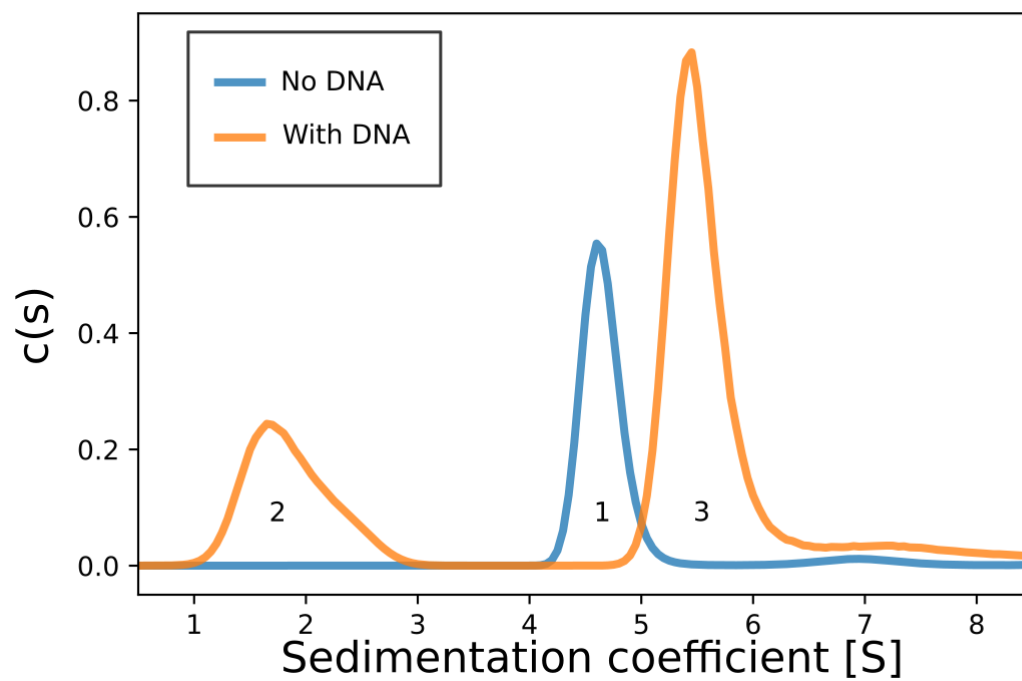


Figure S2. Comparison of BamHI, EcoRI and BsmI scaffolds. The upper panel shows the BamHI homodimer, the EcoRI homodimer and the BsmI monomer. The invariant β -sheets are colored in light blue, the invariant flanking α -helices are colored in red. The middle panel represents only the domains with the catalytic residues. The lower panel is a different orientation (rotation along the vertical axis).

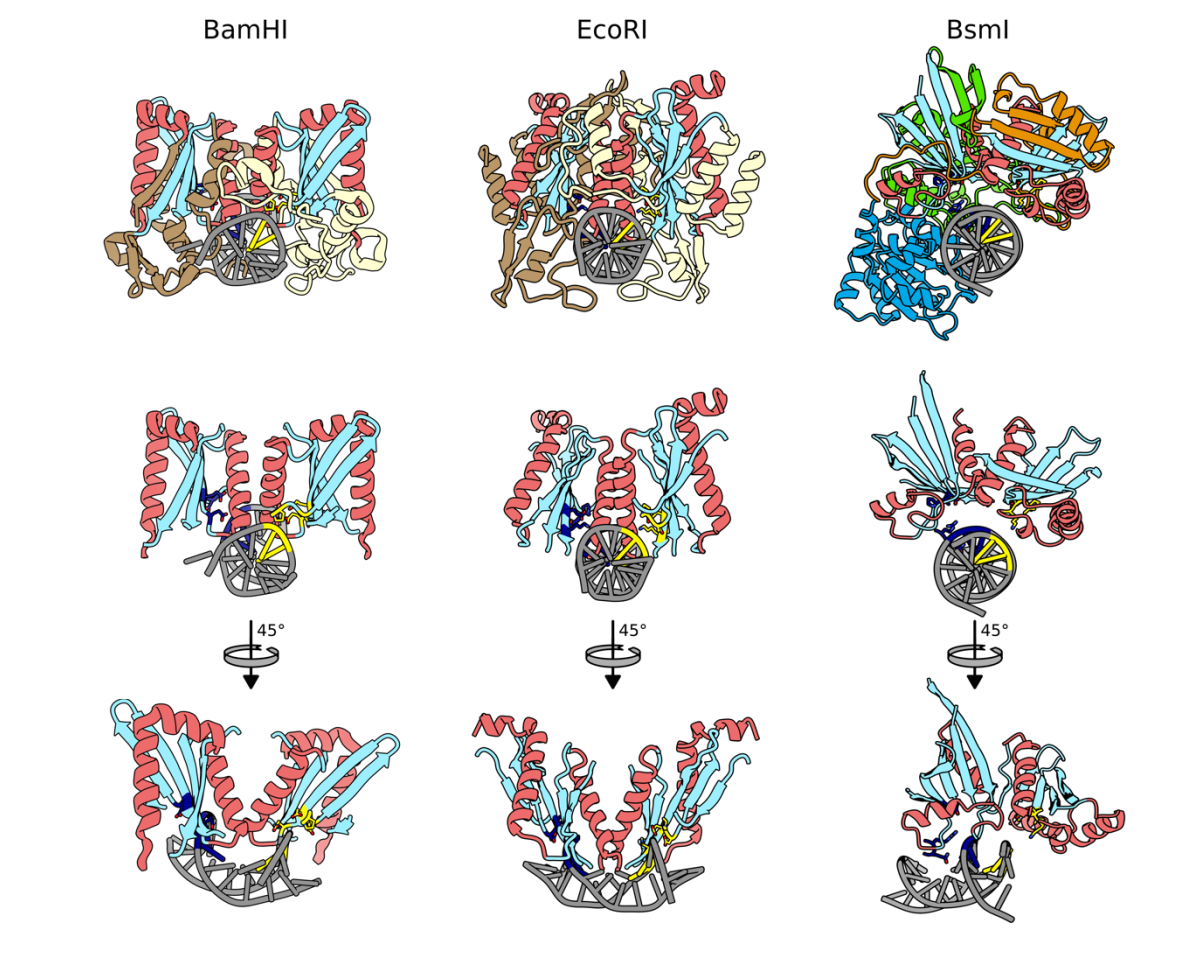


Figure S3. Sequence logo obtained from the MAFFT multi-alignment of the 184 hits from BlastP with a coverage larger than 80 %. Generated with the Weblogo 3 web application (3).

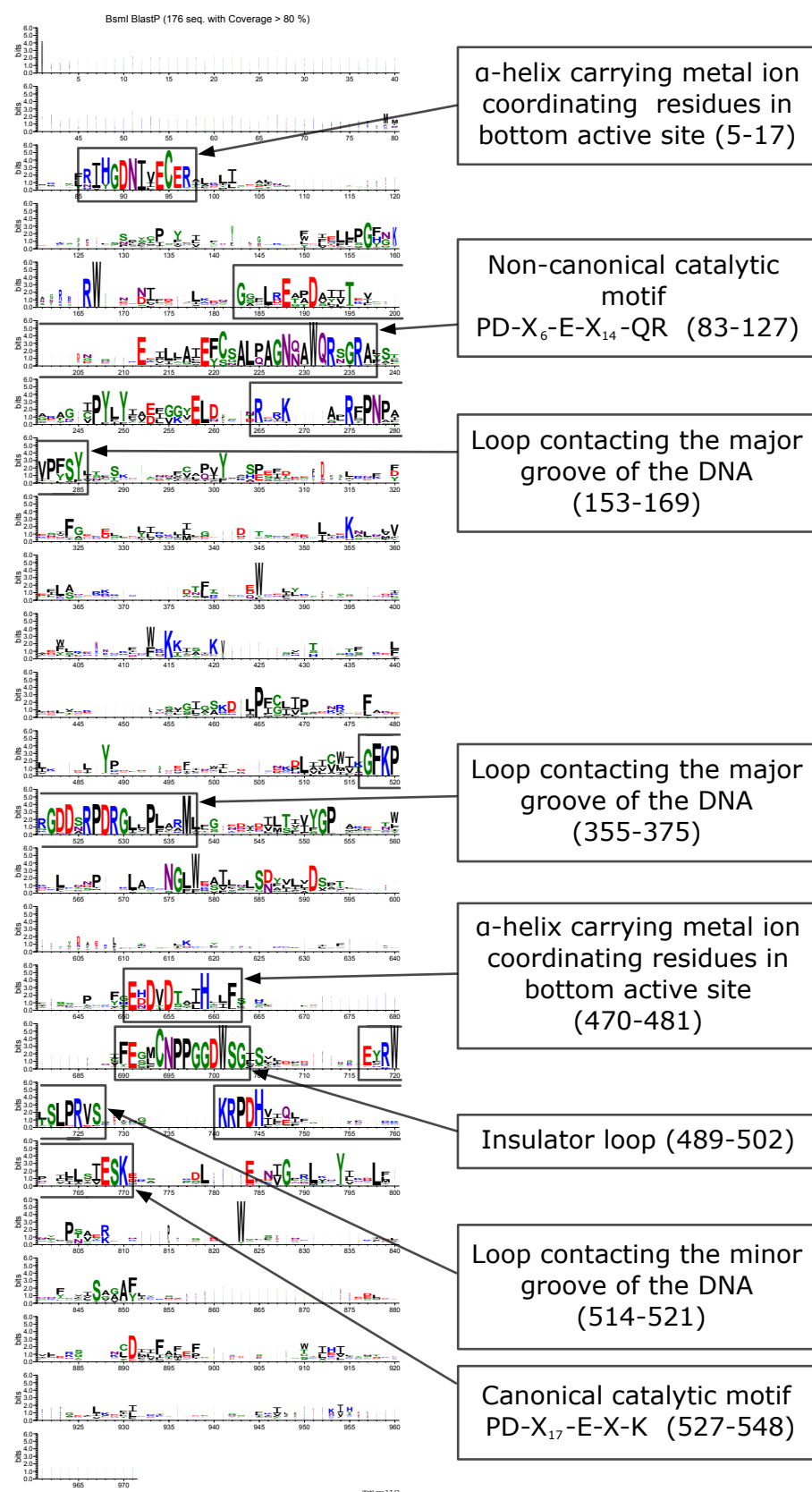
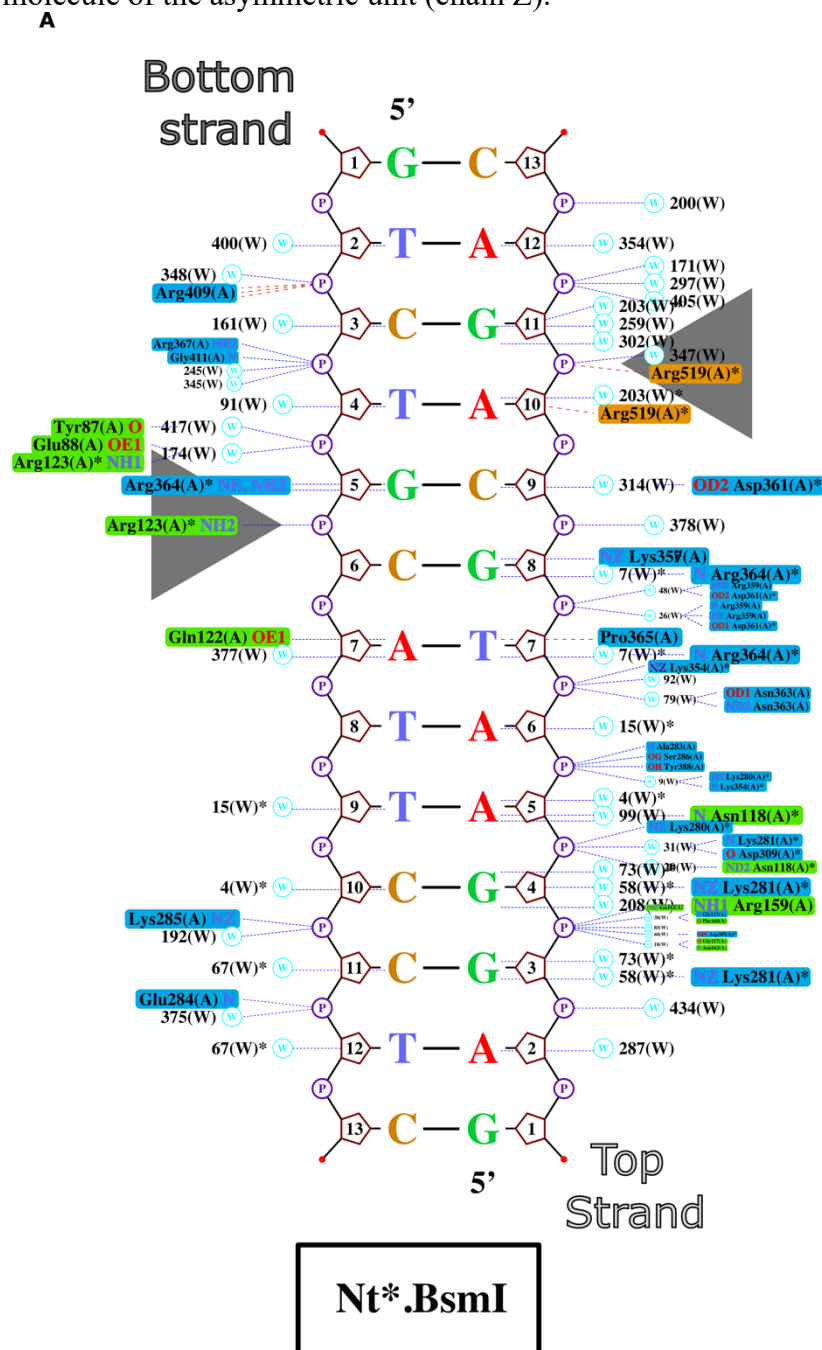


Figure S4. Schematic representation of Protein-DNA interactions for wt.BsmI and Nt*.BsmI. Boxed residues are colored according to the domain the residues belong to. The dark grey triangles indicate the scissile phosphodiester bonds. The residues Asp91, Glu 109, Glu470 and Asp 499 are coordinated by Ca^{2+} ions in the catalytic sites. Adapted from Nucplot results obtained with default parameters (4). (A) Nt*.BsmI (PDB id. 9EZ7). (B) wt.BsmI (PDB id. 9F38), first molecule of the asymmetric unit (chain A). (C) wt.BsmI (PDB id. 9F38), second molecule of the asymmetric unit (chain Z).

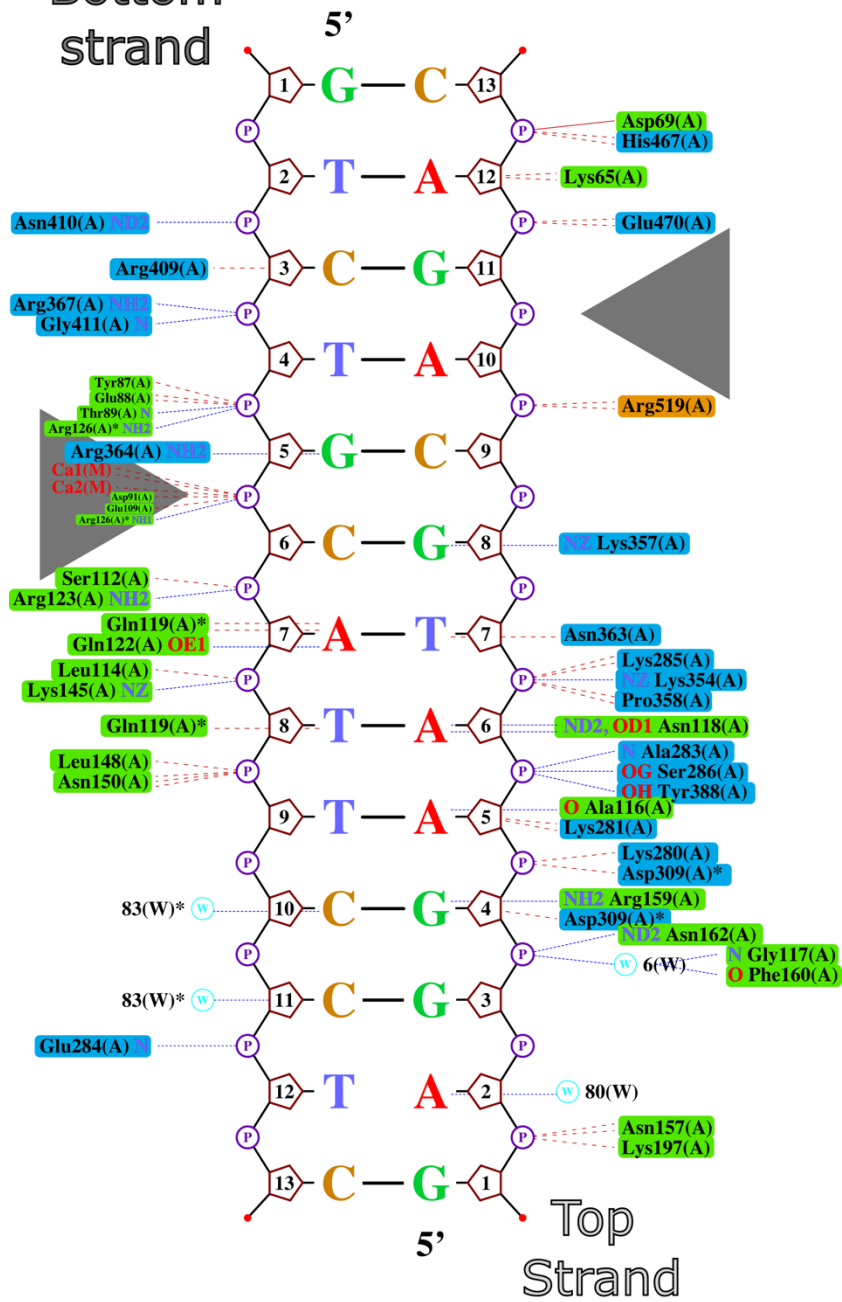


Key

- 3 Backbone sugar and base-number
- P Phosphate group
- * Residue/water on plot more than once
- Hydrogen bond to DNA
- - - Nonbonded contact to DNA (< 3.35Å)
- 88 W Water molecule and number

B

Bottom
strand



wt.BsmI
(chains A, E, F)

c

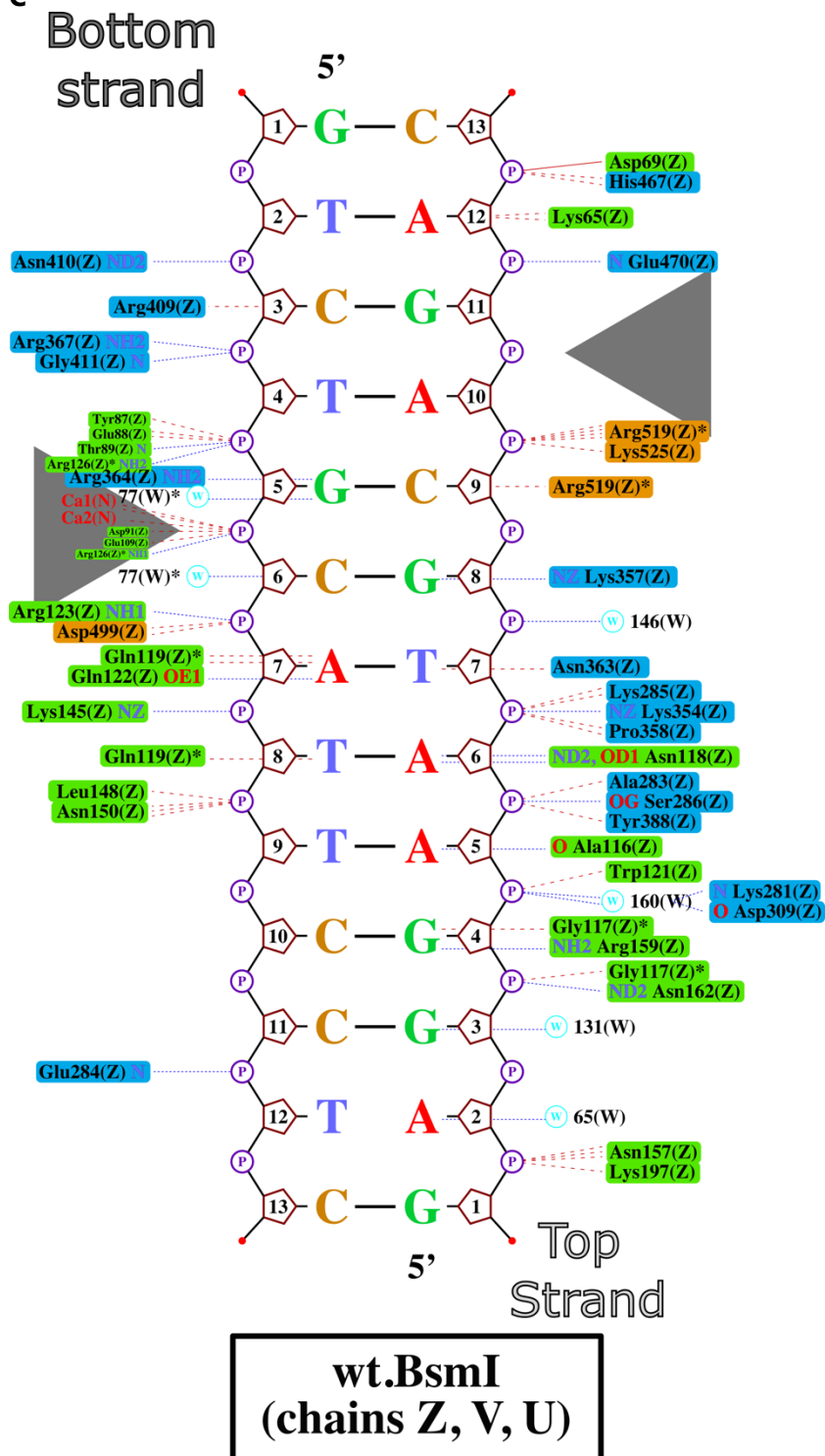


Figure S5. X-ray crystallographic model and AlphaFold3 model comparison. (A) Structural alignment of the apo AlphaFold3 model (dark grey) with the BsmI crystallographic model (PDB 9F38). The alignment and the RMSD were obtained using the Matchmaker tool of ChimeraX (2). (B) X-ray crystallographic model of BsmI. The blue DNA strand corresponds to the bottom strand, the red one to the top strand. The pink balls are Ca^{2+} ions. (C) AlphaFold3 (5) model of BsmI protein with three Ca^{2+} ions and the same cognate dsDNA as the one used for crystallogenesis. The blue DNA strand corresponds to the bottom strand, the red one to the top strand. The pink balls are Ca^{2+} ions.

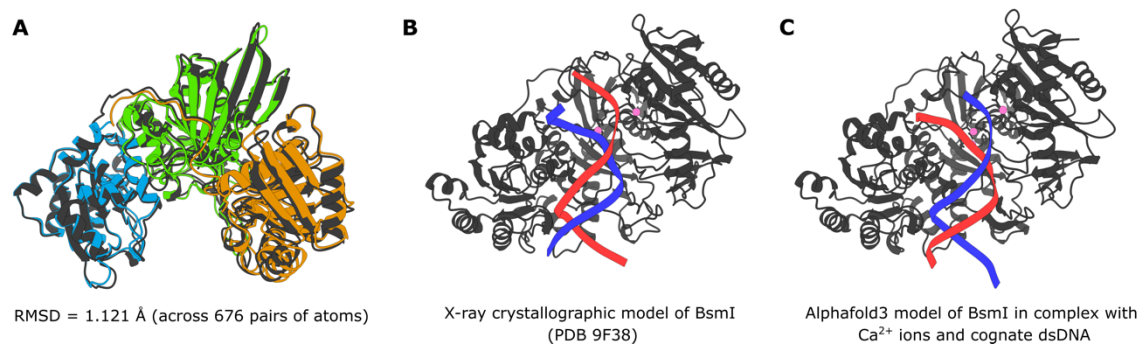
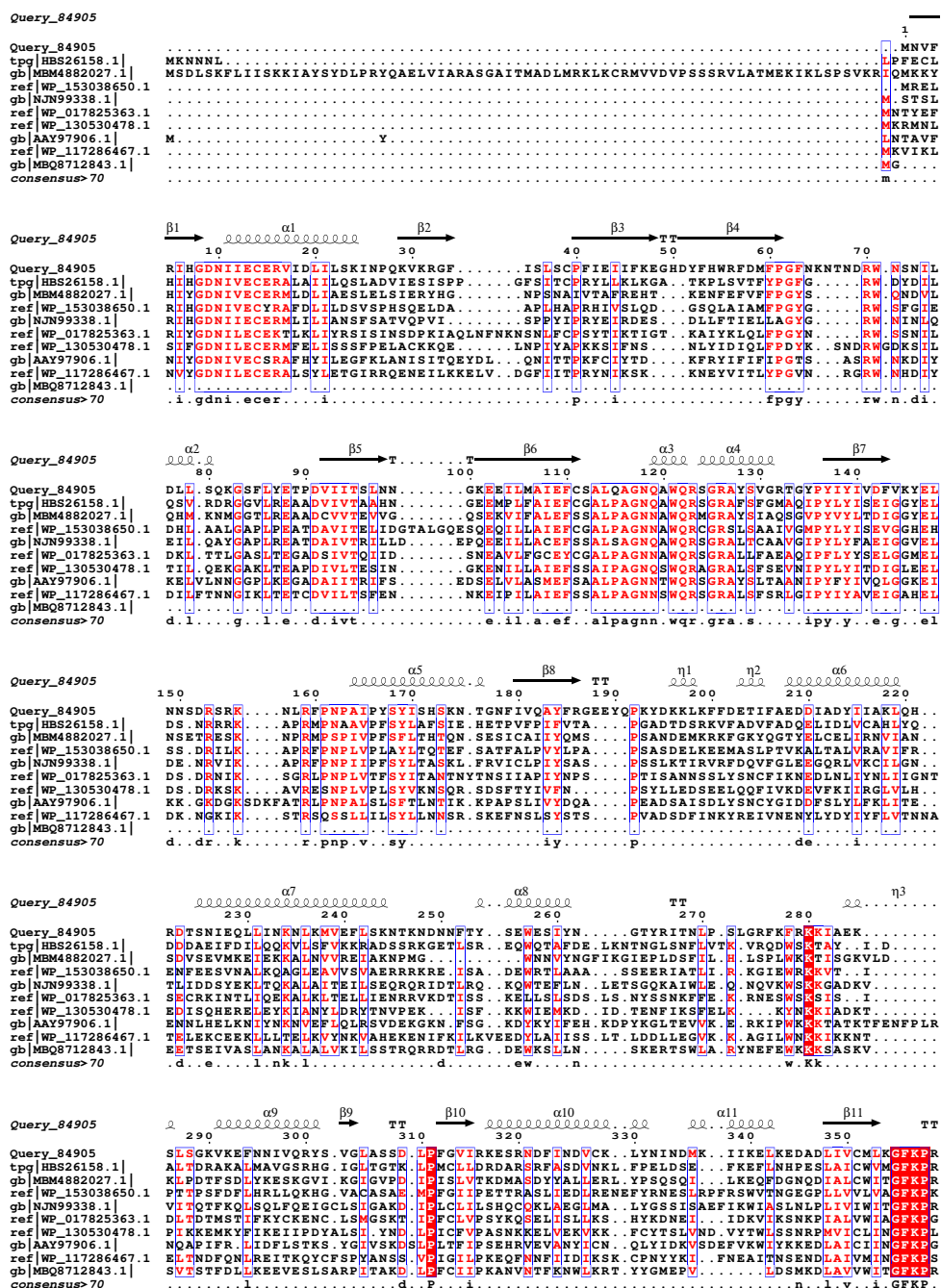


Figure S6. MAFFT Multi-alignment of the 10 sequences presenting less than 40% of identity between each other among the 184 results of BlastP with a coverage larger than 80 %. The gb|AAY97906.1 entry corresponds to Mva1269I. The multi-alignment was then treated with the ESPrpt 3.0 web application (6) to highlight the conserved residues (boxed in red).



Query_84905

360 370 380 390 400 410 420 430

η4 α12 β12 α13 α14 α15 β13 TT

Query_84905

tpg|HBS26158.1|

gb|MM4882027.1|

ref|WP_153038650.1|

gb|JUN99338.1|

ref|WP_017825363.1|

ref|WP_130530478.1|

gb|AA97906.1|

ref|WP_117286467.1|

gb|MBQ8712843.1|

consensus>70

GDDNRPRDGLPLVAMLAGENAQIFTFIYGF...IKGAINLTDQDINKLAKRNGLWKSFVSLSDFIIVLDCPIIGESYNEFR

GDDARPRDGLAPITRMVSSDHEMLTIVYGF...APEATWKAIVKQPEKLAARNGLWEAIMESDALLVEAATDHSVSKRGYI

GDDSRPRDGLPLPLRLVGENVKVISIVYGF...GSVVMWKKFTQSFQSLAERNGLWEIVITNLSDDLADSLNNNEPVIAPT

GDDSRPRDGLPLPLVPMFPCDDVILTIVYGF...GSLASTKRAKETFEQAVRANGLWASVLTLSDRITVDAVNTPLPMDIDS

GDDSRPRDGLPLPLARMFCNEVNIISVYVGF...AKPAMWTMLQNSPQLIRQNGLEAIAIANSDAILVDSATAADNPFALL

GDDSRPRDGLPLPLFKMFCCKEYTVISLVFGF...MKKLAIKTLELTFYNLANSNGLWKAITYTQSDFLIVDNIYISKSKALGIS

GDDSRPRDGLPLPLARMFCNNIDILSLVFGF...ATEIMQHNYKNNPFILASKNGLWKSILSYSSLTADSIHWSLSQSDIY

GDDSRPRDGLPLPLTKMLTNDLILTLMFGF...APPTQWDYLDSDPEKLNKTNGLWQSIFAFSDAILVDSSTRDNKKFVY

GDDSRPRDGLPLPLSRMVFGEEISIFTIISYGF...VSEYIINAIQDGKFDVLRNGLFNAILDGDSDYTLVDSFKHDKCLEIPS

GDDSRPRDGLPLPLCRMLLCKDANILAIYSGF...GTQYTNWLEENFLELCHNSNGLFQAIFVCCNYLVDSITCSKKMFINI

G#DsrPRDGLPLPLRM1.g.#.di...vygfp...e.p...1...NGLW...1...sd...lv#

Query_84905

440 450 460 470 480 490

β14 α16 α17 β15 T

Query_84905

tpg|HBS26158.1|

gb|MM4882027.1|

ref|WP_153038650.1|

gb|JUN99338.1|

ref|WP_017825363.1|

ref|WP_130530478.1|

gb|AA97906.1|

ref|WP_117286467.1|

gb|MBQ8712843.1|

consensus>70

L...IINKNNKESILRKTSKQNNILVDPT.PNHVQENDDVTVIYSIFKYIV...PNCFSQMCNPP

R...SEWVEDLL...EKRSEFFVTPC.PKRIGENDDVTALHLLLSKFIG...KDVFEQMCNPP

G...REVHTGKPEAL...AMKNKELYL.PKKYGENDDVDSIIHIVFKHLLN...DSCFECMCNPP

N...RVSAAPFRDSE...VPSLSAASR.NTAYGENDDVSVIKQLVNS...G...GAFECMCNPP

L...QRSHSHFQGQI...HFTSASA.ITVFTENDDVTVLHLLFAHNHS...RRVFEAMCNPP

NIFDNIYISINPKDIL...IYDTCST.IPRIGENDDVTVIHSIFKTISPFSLANKNNKDSNCPFF

Q...FQLNSPKSSS...KHSISFPFPLKPKKK...DDTAHLTFSSQ...NYIFEELCNPP

Q...NAYLKEHVVQREKKESITPISYF.PKSVGENDDVTSLSLILFTYI...GKHFEFCMCNPP

Q...HSKIPNEI...QNIETPEINRKPERRIGENDDVTVIHMLFKDI...NYVFESEFTNPP

K...TKISPPQEEVI...TFPIYIEKPEVSFFENDDVCAITHQIMAHHEA...IGVFEFCMCNPP

...e...E.DvDt.ih.f...%e.mcNPP

Query_84905

500 510 520 530 540 550 560

T β16 TT β17 β18 β19 η5 α18

Query_84905

tpg|HBS26158.1|

gb|MM4882027.1|

ref|WP_153038650.1|

gb|JUN99338.1|

ref|WP_017825363.1|

ref|WP_130530478.1|

gb|AA97906.1|

ref|WP_117286467.1|

gb|MBQ8712843.1|

consensus>70

GGDWSGLSID...NVHEFRWLSLPRVS.ENGKRPDHIQILDLF.EKPLLLSIVSEKPKNDL...EPKIGVQLIKYIE

GGDWSGVSILGD.SREMEYRWLSLPRVS.GADMKRPDHIQIFAPTCMPSTILSIESETPKSL...ETKICPRLTAYLH

GGDWSGVSLLMGDN...EYRWLSLPRVSRNSKRPDHIQIFDYQN.DNILLSESDNFLRNL...EKDIPSPQLQCE

GGDWSGISLLOP.TDGALEWLSLPRVSGPHRKRPDHIQVHSSV.SSNALISVSKNRPVDF...ESRSEKRRLE

GGDWSGLSLLNF.QTGEYRWLSLPRVSGTGGKRPDHIQIFQLDD.DRLFLLAIESKNRAFNL...ETNVVHHAAITQ

GGDWSGIEFLDI.NNNVIYRWLSLPRVS.GEDKRPDHIQF...SDKILILIESKLHARDL...EKNIPRLINIVQ

GGDWSGIEFLDI...GGTEHRRWLSLPRIS.NESKRPDHIQFACNN.NTYILILIESKEKLLGLL.SDRREDLQSLIDIVL

GGDWSGVSLLK...NNIEYRWLSMYRVSQDGTKRPDHIQVLVNS.TDTLLILIESKGIKNDLLKSEANVGIGIMINYLK

GGDWSGISID...EQYEWYRWLSLREP.KNVRKPDHIQI...KSDKLVIESKEGYSTLI.DKENVCSKMIATLR

GGDWSGIEFYK...GDEYRWLSLPRVS.EYSKRPDHIQIGNGS.MPLFVILIESKEGRDL...NNINRRLKDIIR

GGDWSGIS.1...eyrW.S\$prvs...KRPDHI.#...1.leSKe...d1...e.nic\$.%.%

Query_84905

570 580 590 600 610

η6 β22 α19

Query_84905

tpg|HBS26158.1|

gb|MM4882027.1|

ref|WP_153038650.1|

gb|JUN99338.1|

ref|WP_017825363.1|

ref|WP_130530478.1|

gb|AA97906.1|

ref|WP_117286467.1|

gb|MBQ8712843.1|

consensus>70

YFDFTSP.SVQR...KIAGGNWFGNKS...LVPNDFILLSAGAFI...DYDN...LTENDYEK

YFDFTSP.SIEK...KHSSDWEHSTRK...VSLGDIYRLLSAGAFIMLG.DPEK...VLD

DEFLTSP.NAERVLVPVAGVYPERKGDVGS...VELKARNNVSAVALILGKDKDK...DVEE

ETIEHSP.TVLR...FEEEPRLAKGVDIVWPDGLANTPVLAVGAYCAN.SAKE...EALH

QVTTSP.TIFR...AANAEMDLWQND...ISLSNLTILSGGAFICWL.GTRD...LED

KFNISF.ICHK...KYPLDNVVLNDNVT...INTDSYDYLSAVAFKTK.STIKNVCDVQSVSDISSSYNDGDDSD

KTEEPS.SAVN...IDGHKKSGTK...PNILVETITAVAFYF.KEE...EYTS

NEMARDY.TAVK...KDGEKNIHQ...MTLDKFLTFSAVALFTTDFDN...EYTS

DLEFEPINAKK...KHGEKEYSGTDEK...IDTSYFNLSSAVAFYVK.NDSE...ITRESCK

DDHSEPTAHR...TSANTEWRFFDGT...LGEVAYSLLTIGAFILYK.NDEE...LRR

L...p...ew...lsa.A%.e...

Query_84905

620 630 640 650 660 670

β23 β24 α20 η7 β25

Query_84905

tpg|HBS26158.1|

gb|MM4882027.1|

ref|WP_153038650.1|

gb|JUN99338.1|

ref|WP_017825363.1|

ref|WP_130530478.1|

gb|AA97906.1|

ref|WP_117286467.1|

gb|MBQ8712843.1|

consensus>70

IFEVTGCDLLIAIKNQNNPQKWVIFKPKNTIAEKLVDNYIKLNFK.SNIFDTGFFH...IEG...

AVQRSKTDLTFFVDFLNSGKSCQIFIYSSTSNGREIAEYIAANVED.SEIIAVKVIN...

ILKSNADICPGLSFLPDG.SVDIFVLNKNKN...EXVEKIVDS...N...KITNELKINFIKEL...

VFRSHDLDLLAMFSSSG.ACLIRITNSTA.RLAARSPVSAEPLTPNVE.KVDA...FADEIE...

TLARCOLDIFAIIFRSVEQSALLHIKTGARG.EFLIPQIYNLVQOQGGRLLEIEH...

LFELNLSDIIFDIITHSNI.IFIDIACRNNAI.KKWVETMIDYNR...NNFLSSHKVVFVNTISPVNFFV

AEKLQIDIFIAFNLSSE...VIYKHTPKG.KVVLDFKRNKT...IQ...

LIVHNSQTALAFALIEKEKN.SVMHIFTANTVA.YNFAEYLLLETMR.NSHLPKIKYK...PI...

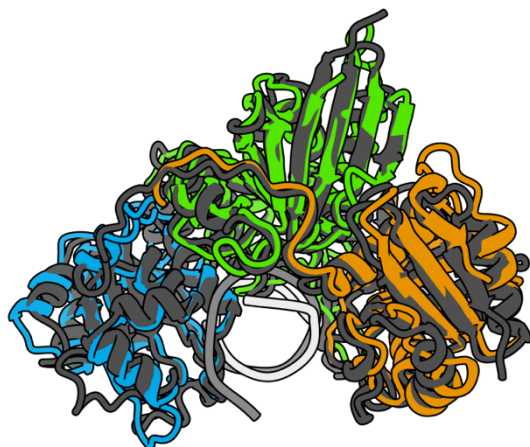
LATNCDVDLVFGIKVTENKHIVIIYYLNRQSN.FNELNLTKDAISLFKNYNEISFKR...VK...

HLDRGNIDALAFEFGEVS...TLHFLSNENG.ALIGDILQIASE.QGSFVVKH...

...n.#...e...

Figure S7. Structural alignment of BsmI and its commercial isoschizomer Mva1269I (1). The Mva1269I structure is a model generated by AlphaFold (v2.1.1) and appears in dark grey. BsmI is colored according to the same color code for domains as in Figure 1. The structural alignment was performed with the ChimeraX (v1.5) ‘matchmaker’ command (2).

Mva1269I vs BsmI



Blastp:

Query Cover = 97%

E-value = 1e-80

Identity = 30.2%

ChimeraX Matchmaker:

RMSD between 388 pruned
atom pairs = 1.275Å

(across all 644 pairs = 3.860Å)

Figure S8. Nuclease activity tests using urea-acrylamide gel electrophoresis. (A) Comparison of the activities of N0.BsmI, Nt.BsmI, Nb.BsmI and BsmI on the three substrates P1, P2 and P3 in the three commercial buffers r2.1, r3.1 and rCutsmart (New England Biolabs, Inc.). The reaction was incubated for 1 hour at 65°C. An incubation of one hour at 37 °C with the proteinase K was performed to guarantee enzyme degradation prior to loading the gel. (B) Control experiment showing the migration of the various fragments composing the probes (“No Enzyme” lanes) and comparison with the activity tests of N0.BsmI, Nt.BsmI, Nb.BsmI and BsmI on the probe P1. “1” is the assembly of the complete precut probe, “2” is the cut bottom strand annealed to its complementary strand, “3” is the cut top strand annealed to its complementary strand, “4” is the top strand alone, “5” is the bottom strand alone, “6” is the cut top strand alone, “7” is the cut bottom strand alone. (C) Scheme of the control strands (1-7) tested on gel (B).

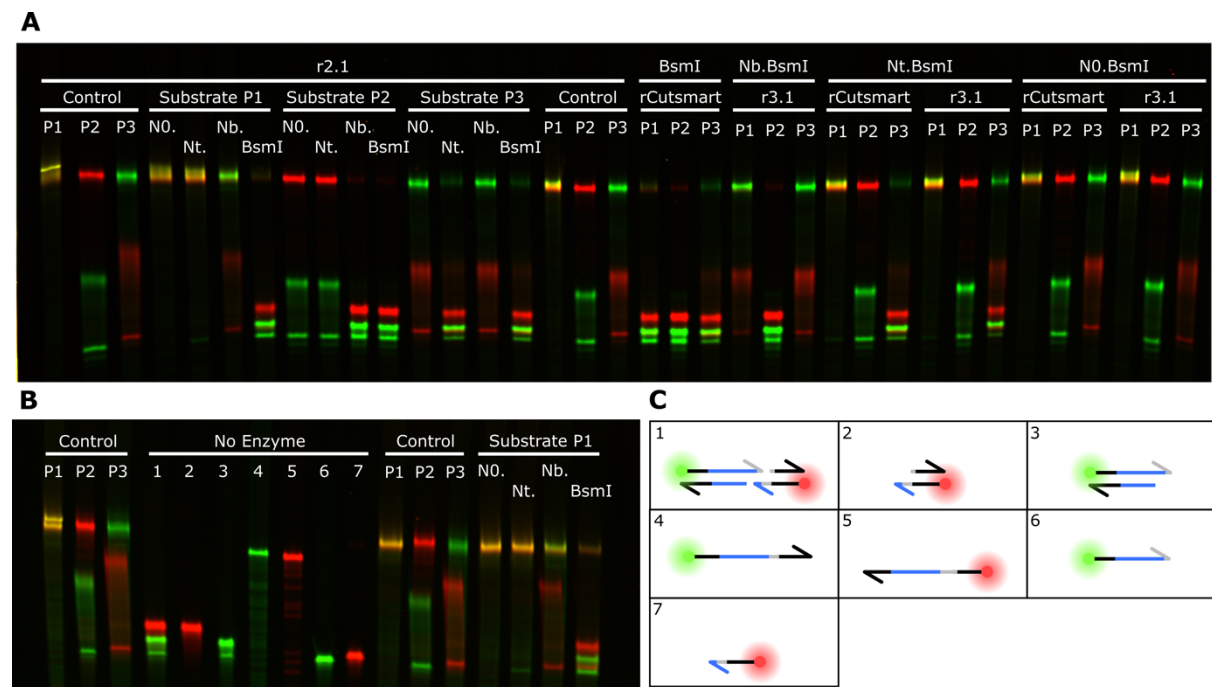
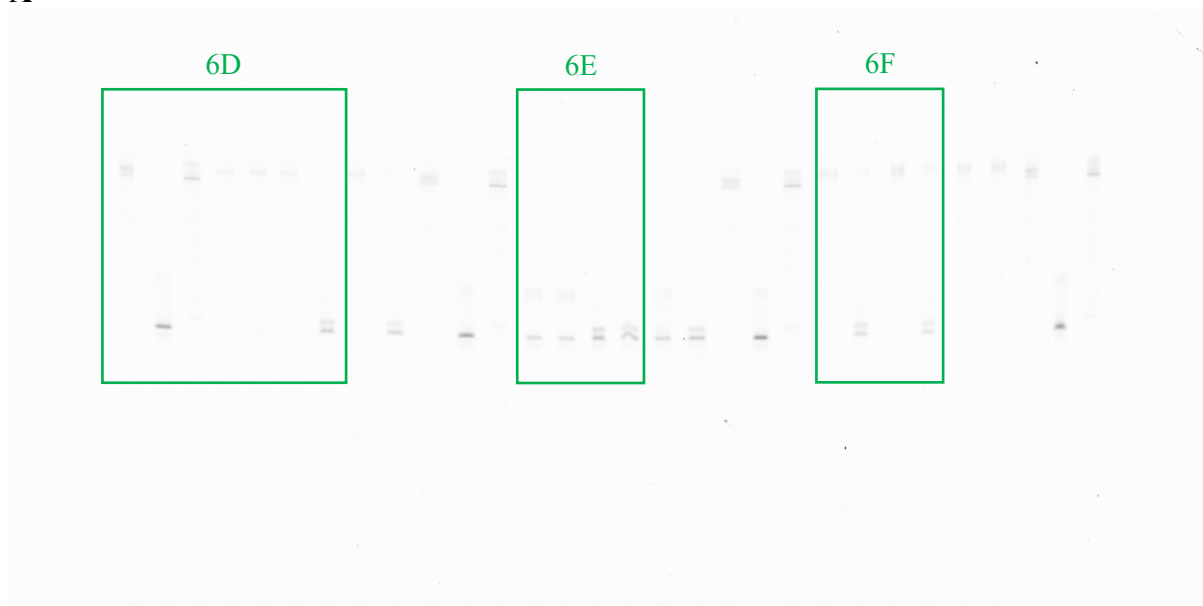


Figure S9. Raw data of the gel electrophoresis acquired in a Typhoon Gel Scanner (GE Healthcare) used to create Figure 6 D-E-F. (A) Acquisition in the FAM channel. (B) Acquisition in the ROX channel.

A



B

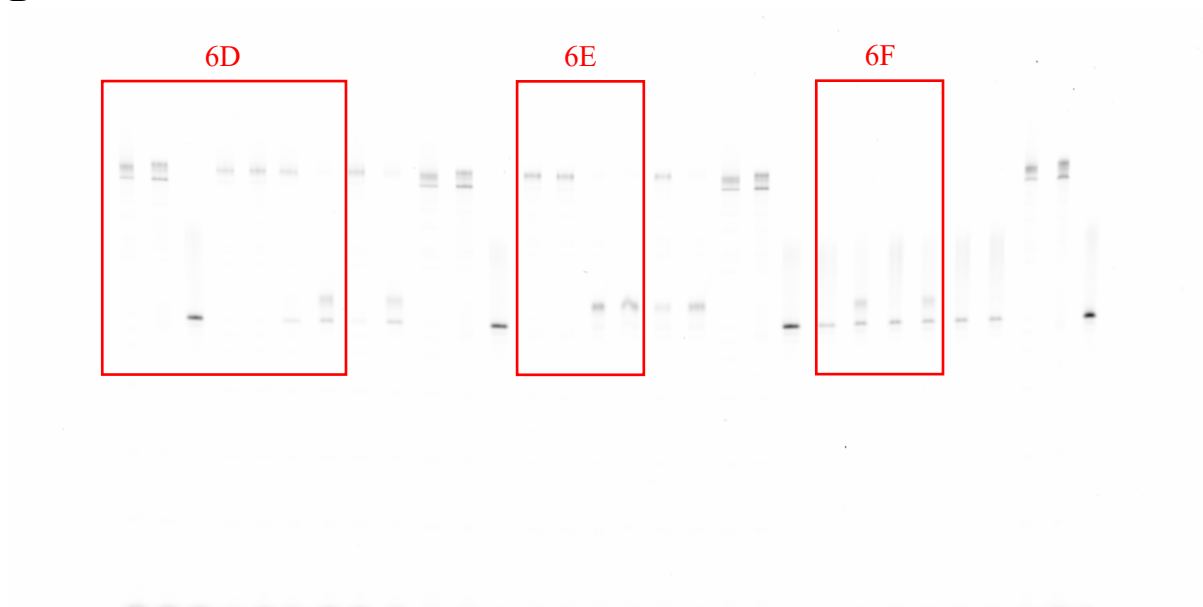


Figure S10. Raw data of the gel electrophoresis acquired in a Typhoon Gel Scanner (GE Healthcare) used to create Figure 7 D-E. (A) Acquisition in the FAM channel. (B) Acquisition in the ROX channel.

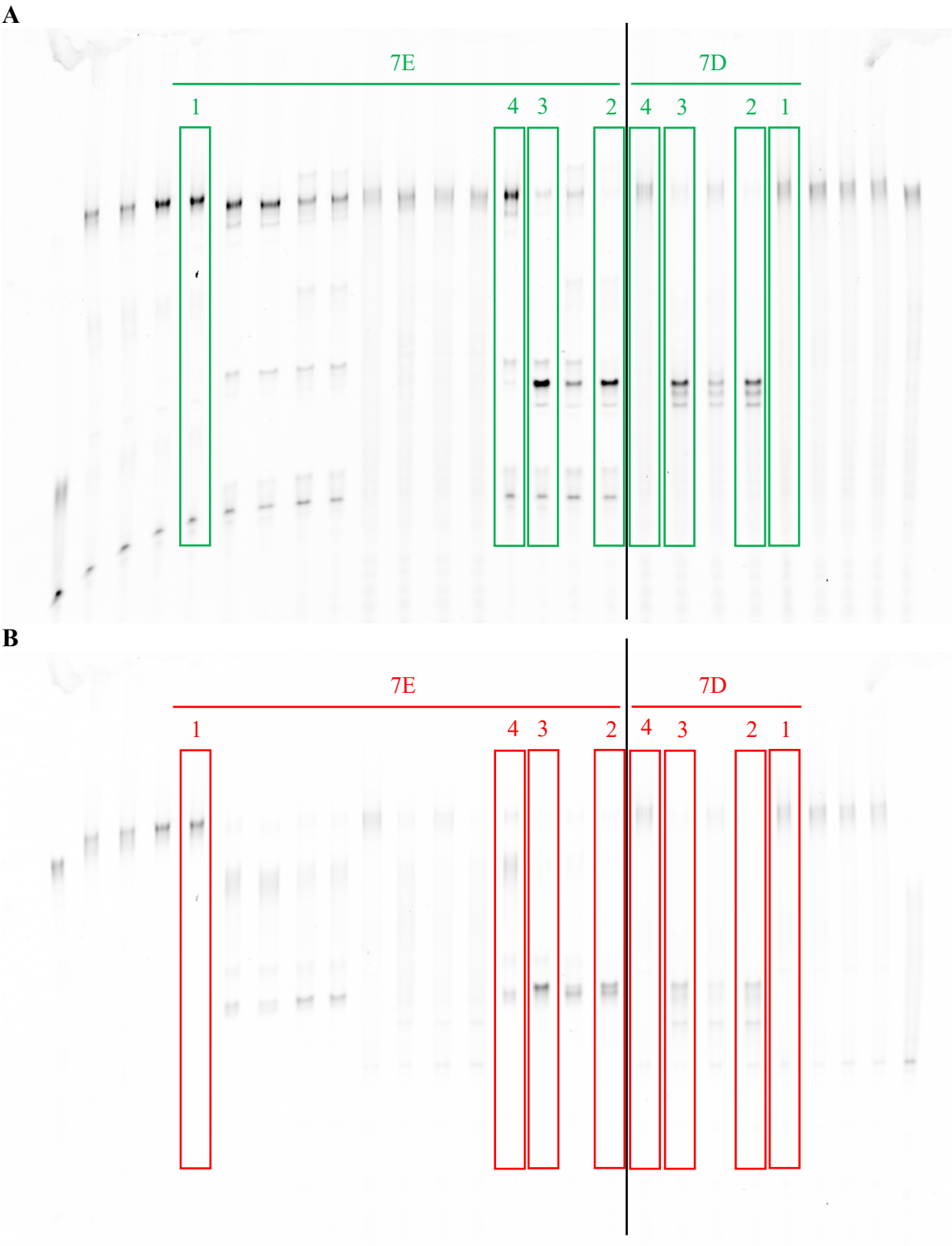


Figure S11. Raw data of the gel electrophoresis acquired in a Typhoon Gel Scanner (GE Healthcare) used to create Figure S8 A. (A) Acquisition in the FAM channel. (B) Acquisition in the ROX channel.

A

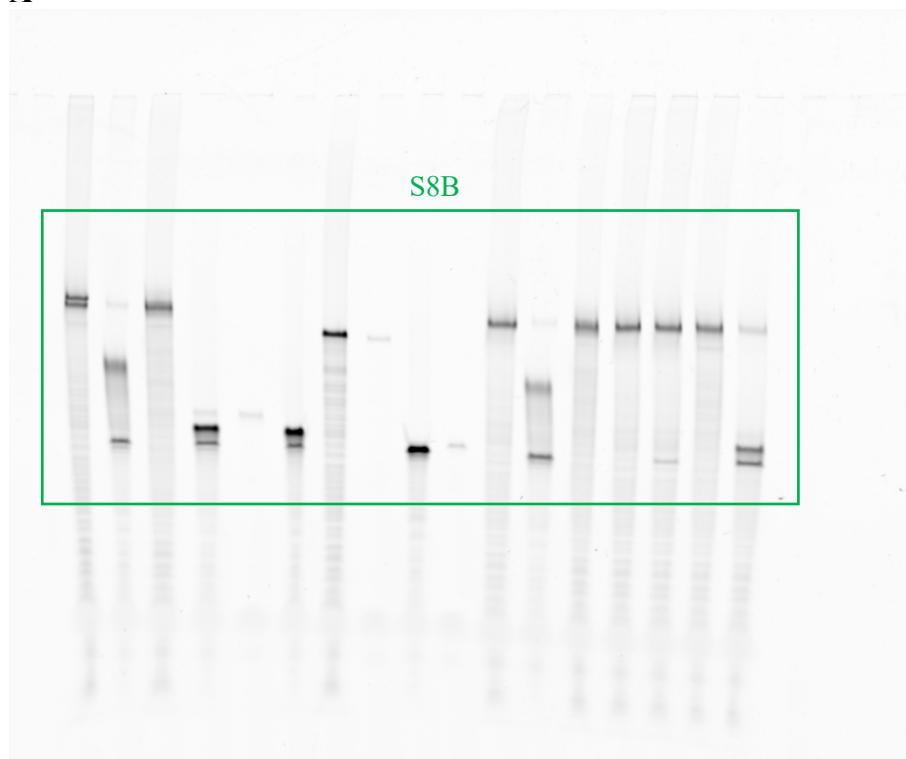


B

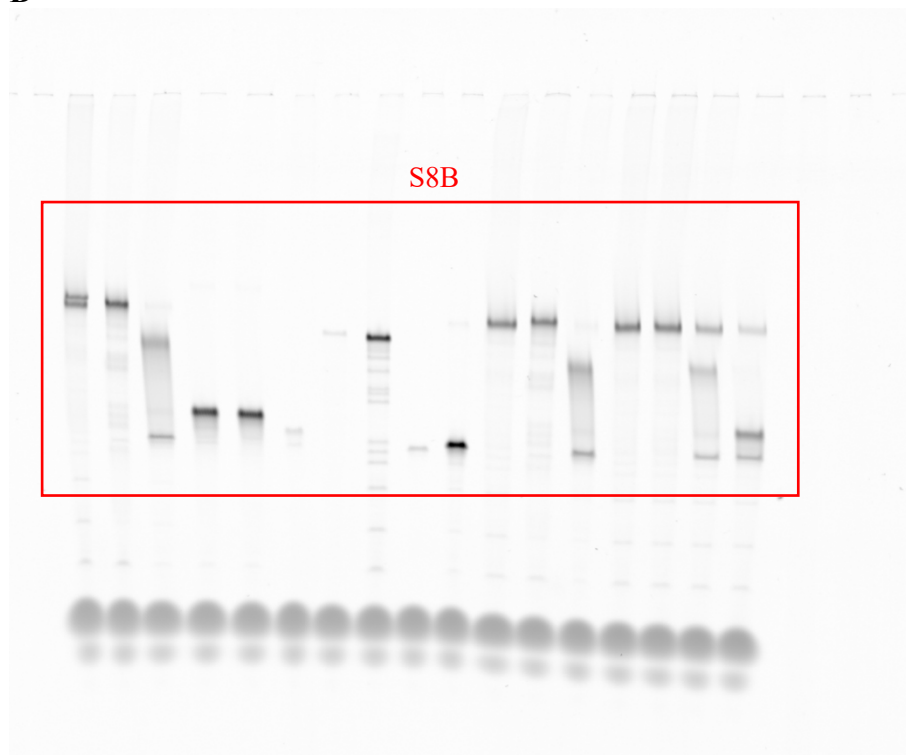


Figure S12. Raw data of the gel electrophoresis acquired in a Typhoon Gel Scanner (GE Healthcare) used to create Figure S8 B. (A) Acquisition in the FAM channel. (B) Acquisition in the ROX channel.

A



B



Supplementary Table S1. List of the 184 results of a Blastp search with BsmI limited down to 80% of coverage and sorted by % of identity.

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. ident	Acc. Len	Accession
1 BsmI [Geobacillus stearothermophilus]	Geobacillus stearothermophilus	1377	1377	100%	0,0E+00	100,00%	676	AA186024.1
2 hypothetical protein [Defluviitaleaceae bacterium]	Defluviitaleaceae bacterium	628	628	84%	0,0E+00	52,67%	583	MCL2572392.1
3 hypothetical protein [Clostridium estertheticum]	Clostridium estertheticum	655	655	94%	0,0E+00	51,07%	700	WP_152749861.1
4 hypothetical protein [Roseburia sp.]	Roseburia sp.	639	639	96%	0,0E+00	49,08%	667	MCM1439624.1
5 hypothetical protein [Sporobacter termitidis]	Sporobacter termitidis	656	656	97%	0,0E+00	48,95%	674	WP_084726616.1
6 hypothetical protein [Dialister invisus]	Dialister invisus	606	606	89%	0,0E+00	48,38%	683	WP_303018806.1
7 hypothetical protein [Blautia sp. LMA6:89]	Blautia sp. LMA6:89	598	598	98%	0,0E+00	47,23%	678	WP_287084107.1
8 hypothetical protein [Lachnospiraceae bacterium]	Lachnospiraceae bacterium	594	594	98%	0,0E+00	46,92%	678	MB05238018.1
9 hypothetical protein [Lachnospiraceae bacterium]	Lachnospiraceae bacterium	575	575	98%	0,0E+00	46,92%	678	MCI8863002.1
10 hypothetical protein [Clostridium liquoris]	Clostridium liquoris	585	585	93%	0,0E+00	46,65%	677	WP_106062445.1
11 hypothetical protein [Parasutterella secunda]	Parasutterella secunda	554	554	89%	0,0E+00	46,54%	683	WP_25834005.1
12 hypothetical protein [Proteiniborus ethanoligenes]	Proteiniborus ethanoligenes	574	574	93%	0,0E+00	46,48%	648	WP_091733038.1
13 hypothetical protein [Volucribacter amazonae]	Volucribacter amazonae	601	601	95%	0,0E+00	46,30%	683	WP_279572375.1
14 hypothetical protein [Fusobacterium periodonticum]	Fusobacterium periodonticum	572	572	98%	0,0E+00	46,09%	678	WP_005971533.1
15 hypothetical protein [Ligilactobacillus salivarius]	Ligilactobacillus salivarius	552	552	86%	0,0E+00	45,87%	595	WP_180376193.1
16 hypothetical protein FACS1894167_0580 [Synergistales bacterium]	Synergistales bacterium	552	552	94%	0,0E+00	45,78%	684	GHV28538.1
17 hypothetical protein AGMMS49983_09520 [Clostridia bacterium]	Clostridia bacterium	548	548	94%	0,0E+00	45,62%	688	GHU64238.1
18 hypothetical protein [Bacillus altitudinis]	Bacillus altitudinis	587	587	99%	0,0E+00	45,60%	677	MED0851892.1
19 TPA: hypothetical protein [Defluviitaleaceae bacterium]	Defluviitaleaceae bacterium	582	582	99%	0,0E+00	45,60%	677	HPT77230.1
20 hypothetical protein [Acidaminococcus fermentans]	Acidaminococcus fermentans	571	571	97%	0,0E+00	45,60%	680	WP_154488117.1
21 hypothetical protein [Clostridiales bacterium]	Clostridiales bacterium	590	590	97%	0,0E+00	45,45%	680	MD07085157.1
22 hypothetical protein [Collinsella sp. AM20-15AC]	Collinsella sp. AM20-15AC	593	593	98%	0,0E+00	45,41%	677	WP_147357130.1
23 hypothetical protein [Selenomonas sp.]	Selenomonas sp.	572	572	97%	0,0E+00	45,13%	684	MB06204506.1
24 hypothetical protein [Selenomonas ruminantium]	Selenomonas ruminantium	579	579	97%	0,0E+00	45,04%	684	WP_303840814.1
25 hypothetical protein [Gottschalkia purinilytica]	Gottschalkia purinilytica	578	578	99%	0,0E+00	44,90%	677	WP_050354693.1
26 hypothetical protein [Haemophilus haemolyticus]	Haemophilus haemolyticus	573	573	96%	0,0E+00	44,84%	680	WP_118796360.1
27 hypothetical protein [Mergibacter septicus]	Mergibacter septicus	572	572	97%	0,0E+00	44,76%	680	WP_261920681.1
28 hypothetical protein [Holophagales bacterium]	Holophagales bacterium	530	530	96%	2,0E-176	44,74%	686	MDR2560808.1
29 hypothetical protein [Pectinatus frisingensis]	Pectinatus frisingensis	590	590	97%	0,0E+00	44,68%	677	WP_196060698.1
30 TPA: hypothetical protein [Sedimentibacter sp.]	Sedimentibacter sp.	587	587	97%	0,0E+00	44,68%	677	HOW23787.1
31 hypothetical protein [Flavonifractor sp. An10]	Flavonifractor sp. An10	571	571	97%	0,0E+00	44,61%	679	WP_087237546.1
32 hypothetical protein [Pasteurella multocida]	Pasteurella multocida	569	569	97%	0,0E+00	44,61%	680	WP_324638937.1
33 hypothetical protein [Faecalibacterium gallinarum]	Faecalibacterium gallinarum	573	573	97%	0,0E+00	44,56%	680	WP_238317900.1
34 hypothetical protein [Leptotrichia]	Leptotrichia	576	576	97%	0,0E+00	44,53%	680	WP_018450723.1
35 hypothetical protein [Lachnospiraceae bacterium]	Lachnospiraceae bacterium	580	580	99%	0,0E+00	44,51%	678	MCI9135513.1
36 hypothetical protein [Clostridiales bacterium]	Clostridiales bacterium	565	565	98%	0,0E+00	44,41%	679	MCD7802131.1
37 hypothetical protein [Megasphaera elsdenii]	Megasphaera elsdenii	551	551	97%	0,0E+00	44,41%	680	WP_095340908.1
38 hypothetical protein [Lachnospiraceae bacterium]	Lachnospiraceae bacterium	551	551	98%	0,0E+00	44,23%	678	MCI8527668.1
39 hypothetical protein [Lachnospiraceae bacterium]	Lachnospiraceae bacterium	582	582	99%	0,0E+00	44,22%	678	MCI8997104.1
40 hypothetical protein [Lachnospiraceae bacterium]	Lachnospiraceae bacterium	582	582	99%	0,0E+00	44,22%	678	MCI8854294.1
41 hypothetical protein [unclassified Ruminococcus]	unclassified Ruminococcus	559	559	97%	0,0E+00	44,13%	676	WP_255882670.1
42 hypothetical protein [Ruminococcus sp. zg-924]	Ruminococcus sp. zg-924	559	559	97%	0,0E+00	44,13%	677	MCCQ4022038.1
43 hypothetical protein [Clostridia bacterium]	Clostridia bacterium	559	559	97%	0,0E+00	44,08%	680	MCI8362964.1
44 conserved protein of unknown function [Ruminococcaceae bacterium B	Ruminococcaceae bacterium B	562	562	96%	0,0E+00	44,01%	675	CAB1254145.1
45 hypothetical protein [Sporofaciens muscui]	Sporofaciens muscui	557	557	97%	0,0E+00	43,93%	680	WP_288175471.1
46 hypothetical protein [Sporofaciens muscui]	Sporofaciens muscui	556	556	97%	0,0E+00	43,93%	680	WP_159750721.1
47 hypothetical protein [Candidatus Spyrodocola sp.]	Candidatus Spyrodocola sp.	550	550	90%	0,0E+00	43,93%	636	MDY4518495.1
48 hypothetical protein [Phascolarctobacterium sp.]	Phascolarctobacterium sp.	562	562	96%	0,0E+00	43,84%	680	MB05167960.1
49 hypothetical protein [uncultured Ruminococcus sp.]	uncultured Ruminococcus sp.	578	578	97%	0,0E+00	43,83%	683	WP_302681849.1
50 hypothetical protein [uncultured Blautia sp.]	uncultured Blautia sp.	530	530	97%	2,0E-176	43,70%	683	WP_294678939.1
51 hypothetical protein [Candidatus Gastranaerophilales bacterium]	Candidatus Gastranaerophilales bacterium	566	566	97%	0,0E+00	43,70%	683	MCM1045094.1
52 hypothetical protein [Methanobrevibacter sp.]	Methanobrevibacter sp.	578	578	99%	0,0E+00	43,48%	685	WP_292776643.1
53 hypothetical protein [Trueperella pyogenes]	Trueperella pyogenes	546	546	94%	0,0E+00	43,40%	685	WP_167801479.1
54 hypothetical protein [Oscillospiraceae bacterium]	Oscillospiraceae bacterium	565	565	99%	0,0E+00	43,38%	683	MBQ3542160.1
55 Uncharacterised protein [Mycobacterium tuberculosis]	Mycobacterium tuberculosis	506	506	95%	2,0E-167	43,03%	687	CNI01306.1
56 hypothetical protein [Lachnospiraceae bacterium]	Lachnospiraceae bacterium	555	555	98%	0,0E+00	42,90%	683	MDY4969588.1
57 hypothetical protein [Candidatus Delongbacteria bacterium]	Candidatus Delongbacteria bacterium	499	499	93%	8,0E-165	42,81%	677	MBN2835143.1
58 hypothetical protein [Clostridia bacterium]	Clostridia bacterium	530	530	95%	8,0E-177	42,62%	679	MBQ2932250.1
59 bsmI [Clostridium sp. CAG:921]	Clostridium sp. CAG:921	511	511	96%	1,0E-169	42,51%	657	CCY70910.1
60 hypothetical protein [Oscillospiraceae bacterium]	Oscillospiraceae bacterium	531	531	96%	5,0E-177	42,50%	681	MBE6809800.1
61 hypothetical protein [Clostridia bacterium]	Clostridia bacterium	469	469	90%	2,0E-153	41,94%	638	MCI8273141.1
62 hypothetical protein [Rothia mucilaginosa]	Rothia mucilaginosa	502	502	94%	1,0E-165	41,87%	687	WP_152901967.1
63 hypothetical protein [Rothia dentocariosa]	Rothia dentocariosa	501	501	94%	2,0E-165	41,72%	687	WP_278860064.1
64 hypothetical protein [Corynebacterium sp. H130]	Corynebacterium sp. H130	461	461	90%	6,0E-150	40,26%	655	WP_339985683.1
65 hypothetical protein QEU98_06790 [Trueperella pyogenes]	Trueperella pyogenes	456	456	89%	5,0E-148	40,26%	655	WZ155820.1
66 hypothetical protein [Lentisphaeria bacterium]	Lentisphaeria bacterium	395	395	93%	3,0E-124	37,81%	656	MBR6240216.1
67 hypothetical protein [Spirochaetia bacterium]	Spirochaetia bacterium	376	376	86%	4,0E-118	36,73%	577	MCI7110039.1
68 hypothetical protein [Treponema sp.]	Treponema sp.	400	400	92%	1,0E-126	36,42%	658	MBS7265853.1
69 hypothetical protein [Spirochaetia bacterium]	Spirochaetia bacterium	399	399	92%	8,0E-126	36,38%	659	MCI7564521.1
70 hypothetical protein [Candidatus Micrarchaeota archaeon]	Candidatus Micrarchaeota archaeon	396	396	97%	1,0E-124	36,34%	670	MCT571336.1
71 hypothetical protein [Treponema sp.]	Treponema sp.	394	394	92%	5,0E-124	36,29%	658	MEL3905591.1
72 hypothetical protein [Spirochaetia bacterium]	Spirochaetia bacterium	399	399	92%	8,0E-126	36,15%	658	MCI6808881.1
73 hypothetical protein [Methylophaga aminisulfivorans]	Methylophaga aminisulfivorans	384	384	97%	8,0E-120	36,00%	670	WP_007144371.1
74 hypothetical protein ABL308_06800 [Oceanicaulis sp.]	Oceanicaulis sp.	424	424	98%	2,0E-135	35,99%	663	XBQ17584.1
75 hypothetical protein [Caldilineaceae bacterium]	Caldilineaceae bacterium	389	389	96%	4,0E-122	35,63%	671	MCY3709916.1
76 hypothetical protein [Litoreibacter halocynthiae]	Litoreibacter halocynthiae	384	384	92%	4,0E-120	35,56%	677	WP_281859576.1
77 hypothetical protein [Candidatus Azobacteroides sp.]	Candidatus Azobacteroides sp.	319	319	89%	2,0E-95	35,55%	648	MCL2651293.1
78 TPA: hypothetical protein [Chloroflexota bacterium]	Chloroflexota bacterium	327	327	89%	5,0E-98	35,26%	666	HEY85950.1
79 hypothetical protein NO2_0358 [Candidatus Termitenax persephoneae]	Candidatus Termitenax persephoneae	341	341	85%	1,0E-103	35,15%	660	GBR75708.1
80 hypothetical protein [Treponema sp.]	Treponema sp.	381	381	92%	4,0E-119	35,12%	658	MDY2844220.1
81 hypothetical protein [Candidatus Poribacteria bacterium]	Candidatus Poribacteria bacterium	330	330	93%	2,0E-99	34,98%	668	MF C1714365.1
82 hypothetical protein [Candidatus Omnitrophota bacterium]	Candidatus Omnitrophota bacterium	284	284	80%	1,0E-82	34,95%	584	MBF0490877.1
83 hypothetical protein [Bacteroidota bacterium]	Bacteroidota bacterium	305	305	90%	6,0E-90	34,90%	661	MCB0514788.1
84 hypothetical protein [Bacteroidota bacterium]	Bacteroidota bacterium	313	313	89%	4,0E-93	34,88%	661	MBS1570818.1
85 hypothetical protein E6Q89_00010 [Bacteroidia bacterium]	Bacteroidia bacterium	313	313	89%	5,0E-93	34,88%	661	TXH60935.1
86 TPA: hypothetical protein [Sedimentisphaerales bacterium]	Sedimentisphaerales bacterium	407	407	97%	5,0E-129	34,82%	666	HSV27330.1
87 hypothetical protein [Bacteroidota bacterium]	Bacteroidota bacterium	300	300	90%	3,0E-88	34,76%	625	MEI6435443.1
88 TPA: hypothetical protein [Bacteroidia bacterium]	Bacteroidia bacterium	299	299	90%	1,0E-87	34,59%	661	HLG35431.1
89 hypothetical protein [Vibrio parahaemolyticus]	Vibrio parahaemolyticus	345	345	85%	3,0E-104	34,45%	739	ENM1126097.1
90 hypothetical protein [Candidatus Babeliaceae bacterium]	Candidatus Babeliaceae bacterium	380	380	94%	2,0E-118	34,38%	666	MBN2267646.1
91 hypothetical protein [unclassified Staphylococcus]	unclassified Staphylococcus	285	285	82%	4,0E-83	34,36%	599	WP_210128570.1
92 hypothetical protein [Bacteroidota bacterium]	Bacteroidota bacterium	311	311	93%	2,0E-92	34,35%	661	MEN9448876.1

Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per. ident	Acc. Len	Accession
93 hypothetical protein [Bacteroidota bacterium]	Bacteroidota bacterium	305	305	93%	4,0E-90	34,29%	661	MB51572830.1
94 hypothetical protein A2275_04880 [Bacteroidetes bacterium RIFOXYA1]	Bacteroidetes bacterium RIFOX	308	308	90%	5,0E-91	34,28%	662	OFY72705.1
95 hypothetical protein [Vibrio parahaemolyticus O1]	Vibrio parahaemolyticus O1	342	342	85%	6,0E-103	34,28%	739	EG0942128.1
96 hypothetical protein [Vibrio parahaemolyticus]	Vibrio parahaemolyticus	341	341	85%	7,0E-103	34,28%	739	EG09416292.1
97 hypothetical protein [Vibrio parahaemolyticus]	Vibrio parahaemolyticus	341	341	85%	8,0E-103	34,28%	739	WP_274028301.1
98 hypothetical protein BW200_00308 [Bacteroidetes bacterium ADurb.Bi]	Bacteroidetes bacterium ADurb	311	311	93%	2,0E-92	34,27%	651	OPZ31915.1
99 hypothetical protein [Hugenholtzia roseola]	Hugenholtzia roseola	309	309	91%	1,0E-91	34,23%	641	WP_027001349.1
100 hypothetical protein QR26_17445 [Clostridium baratii]	Clostridium baratii	305	305	92%	2,0E-89	34,22%	691	WWU64762.1
101 hypothetical protein [Flavobacterium sp.]	Flavobacterium sp.	308	308	90%	5,0E-91	34,22%	661	MDI9364355.1
102 hypothetical protein [Bacteroidales bacterium]	Bacteroidales bacterium	297	297	90%	5,0E-87	34,19%	643	MB07083512.1
103 hypothetical protein [Prevotellaceae bacterium]	Prevotellaceae bacterium	313	313	93%	3,0E-93	34,16%	648	MDR2651817.1
104 hypothetical protein [Treponema sp.]	Treponema sp.	298	298	90%	3,0E-87	34,16%	672	MDR3131152.1
105 TPA: hypothetical protein [Salinivirgaceae bacterium]	Salinivirgaceae bacterium	315	315	93%	7,0E-94	34,10%	661	HPW66994.1
106 TPA: hypothetical protein [Gammaproteobacteria bacterium]	Gammaproteobacteria bacterium	386	386	98%	6,0E-121	34,07%	673	HBS26158.1
107 hypothetical protein [Saprospiraceae bacterium]	Saprospiraceae bacterium	303	303	89%	2,0E-89	34,03%	661	MBK9213450.1
108 TPA: hypothetical protein [Rhizomicrobium sp.]	Rhizomicrobium sp.	370	370	96%	9,0E-115	33,98%	666	HEX3809194.1
109 TPA: hypothetical protein [Candidatus Pacearchaeota archaeon]	Candidatus Pacearchaeota arch	306	306	89%	3,0E-90	33,98%	666	H0X29779.1
110 hypothetical protein [Chryseobacterium oryctis]	Chryseobacterium oryctis	308	308	93%	4,0E-91	33,89%	661	WP_264742117.1
111 hypothetical protein [Saprospiraceae bacterium]	Saprospiraceae bacterium	318	318	97%	5,0E-95	33,78%	660	MDZ7879858.1
112 hypothetical protein [Jeotgalibacillus alimentarius]	Jeotgalibacillus alimentarius	271	271	89%	6,0E-77	33,75%	676	WP_041123950.1
113 hypothetical protein [Bacteroidales bacterium]	Bacteroidales bacterium	308	308	93%	4,0E-91	33,70%	648	MDR1974464.1
114 hypothetical protein [Candidatus Bathyarchaeota archaeon]	Candidatus Bathyarchaeota arch	294	294	81%	4,0E-86	33,69%	614	MCJ7767810.1
115 hypothetical protein [Staphylococcus haemolyticus]	Staphylococcus haemolyticus	285	285	83%	2,0E-83	33,67%	568	WP_053027555.1
116 hypothetical protein [Patescibacteria group bacterium]	Patescibacteria group bacterium	292	292	89%	4,0E-85	33,65%	663	MBU447738.1
117 hypothetical protein [Chloroflexota bacterium]	Chloroflexota bacterium	315	315	89%	1,0E-93	33,60%	668	NMB62300.1
118 hypothetical protein [Bacteroidales bacterium]	Bacteroidales bacterium	308	308	92%	5,0E-91	33,54%	648	MDR2407768.1
119 hypothetical protein [Parcubacteria group bacterium]	Parcubacteria group bacterium	290	290	93%	3,0E-84	33,49%	664	MFA7208746.1
120 hypothetical protein AGMMS49574_26260 [Bacteroidia bacterium]	Bacteroidia bacterium	312	312	93%	1,0E-92	33,44%	648	GHT35854.1
121 hypothetical protein [Alphaproteobacteria bacterium]	Alphaproteobacteria bacterium	310	310	80%	3,0E-93	33,39%	551	MDA8030712.1
122 hypothetical protein [Aliarcobacter butzleri]	Aliarcobacter butzleri	267	267	86%	7,0E-76	33,33%	646	WP_130530478.1
123 hypothetical protein [Muribaculaceae bacterium]	Muribaculaceae bacterium	306	306	90%	1,0E-90	33,23%	641	MDE6452767.1
124 hypothetical protein [Rhodobacter sp. SW2]	Rhodobacter sp. SW2	337	337	85%	5,0E-103	33,22%	584	WP_156790719.1
125 hypothetical protein [Opitales bacterium]	Opitales bacterium	335	335	96%	4,0E-101	33,19%	669	MDP4658664.1
126 hypothetical protein [unclassified Staphylococcus]	unclassified Staphylococcus	294	294	92%	1,0E-85	33,18%	669	WP_210140041.1
127 hypothetical protein COU72_04035 [Parcubacteria group bacterium CG]	Parcubacteria group bacterium	286	286	89%	8,0E-83	33,17%	665	PIR56843.1
128 hypothetical protein [Deinococcus sp. SL84]	Deinococcus sp. SL84	352	352	96%	7,0E-108	33,13%	666	WP_268094873.1
129 hypothetical protein [Mariniphaga sp.]	Mariniphaga sp.	296	296	93%	1,0E-86	33,08%	661	MBW6533413.1
130 hypothetical protein [Phocaeicola vulgatus]	Phocaeicola vulgatus	296	296	89%	9,0E-87	33,06%	646	WP_160217050.1
131 hypothetical protein [Staphylococcus sp. GDY8P71P]	Staphylococcus sp. GDY8P71P	291	291	92%	1,0E-84	33,03%	669	WP_210126938.1
132 Uncharacterised protein [Candidatus Gugararchaeum adminiculabundu]	Candidatus Gugararchaeum adr	314	314	91%	5,0E-93	33,02%	691	VVB65454.1
133 hypothetical protein [Clostridium]	Clostridium	290	290	93%	7,0E-84	32,94%	703	WP_043853764.1
134 hypothetical protein [Patescibacteria group bacterium]	Patescibacteria group bacterium	294	294	93%	1,0E-85	32,92%	666	MBP6945358.1
135 TPA: hypothetical protein [Xanthobacteraceae bacterium]	Xanthobacteraceae bacterium	373	373	98%	1,0E-115	32,89%	667	HU233097.1
136 hypothetical protein [Candidatus Taylorbacteria bacterium]	Candidatus Taylorbacteria bacte	302	302	89%	7,0E-89	32,85%	665	MDP3696470.1
137 hypothetical protein [Bacteroidales bacterium]	Bacteroidales bacterium	300	300	92%	2,0E-88	32,76%	643	MBD5236872.1
138 hypothetical protein [Bacteroidales bacterium]	Bacteroidales bacterium	300	300	88%	2,0E-88	32,74%	643	MBR5910742.1
139 hypothetical protein [Candidatus Equimonas enterica]	Candidatus Equimonas enterica	259	259	80%	6,0E-73	32,73%	626	MBQ0157983.1
140 hypothetical protein [Candidatus Margulisibacteriota bacterium]	Candidatus Margulisibacteriota	284	284	89%	9,0E-82	32,69%	677	MEA3493757.1
141 hypothetical protein [Staphylococcus]	Staphylococcus	297	297	92%	7,0E-87	32,67%	671	WP_002487819.1
142 hypothetical protein [Clostridium sp.]	Clostridium sp.	291	291	93%	2,0E-84	32,46%	703	WP_316161014.1
143 hypothetical protein [Clostridium botulinum]	Clostridium botulinum	290	290	93%	1,0E-83	32,46%	710	WP_017825363.1
144 hypothetical protein [Dietzia]	Dietzia	332	332	93%	4,0E-100	32,42%	676	WP_153038650.1
145 hypothetical protein [Carnobacterium maltaromaticum]	Carnobacterium maltaromaticu	268	268	92%	6,0E-76	32,42%	671	WP_317944763.1
146 hypothetical protein [Candidatus Pelagibacter sp.]	Candidatus Pelagibacter sp.	275	275	84%	2,0E-78	32,37%	653	MDC2998355.1
147 hypothetical protein [Bacteroides sp.]	Bacteroides sp.	291	291	93%	1,0E-84	32,21%	641	MBD5353393.1
148 hypothetical protein [Flavobacteriales bacterium]	Flavobacteriales bacterium	275	275	84%	9,0E-80	32,21%	563	MBK6755378.1
149 hypothetical protein [Streptococcus gordonii]	Streptococcus gordonii	231	231	88%	1,0E-62	32,15%	603	WP_227955724.1
150 hypothetical protein [Anaerolineae bacterium]	Anaerolineae bacterium	295	295	94%	4,0E-86	32,12%	662	MGK6625724.1
151 hypothetical protein Broiet011_55400 [Chloroflexota bacterium]	Chloroflexota bacterium	295	295	94%	5,0E-86	32,02%	662	GK141707.1
152 hypothetical protein [Bacteroidales bacterium]	Bacteroidales bacterium	295	295	90%	5,0E-86	32,01%	659	MBQ3614285.1
153 hypothetical protein [Cytophagales bacterium]	Cytophagales bacterium	292	292	89%	4,0E-85	32,01%	658	MBW8050073.1
154 hypothetical protein [Candidatus Omnitrophota bacterium]	Candidatus Omnitrophota bacte	309	309	89%	2,0E-91	31,99%	664	MBU1863902.1
155 hypothetical protein [Muribaculaceae bacterium]	Muribaculaceae bacterium	264	264	81%	4,0E-75	31,97%	602	MDE6019112.1
156 hypothetical protein [Alphaproteobacteria bacterium]	Alphaproteobacteria bacterium	343	343	96%	3,0E-104	31,93%	669	MDA8000226.1
157 hypothetical protein [Solibacillus sp. FSL W7-1464]	Solibacillus sp. FSL W7-1464	268	268	93%	7,0E-76	31,82%	685	WP_340803101.1
158 hypothetical protein [Bacteroides sp. ET71]	Bacteroides sp. ET71	254	254	80%	3,0E-71	31,82%	624	WP_249053436.1
159 hypothetical protein [Anaerolineales bacterium]	Anaerolineales bacterium	301	301	94%	3,0E-88	31,81%	662	NJN99338.1
160 hypothetical protein [Streptococcus intermedius]	Streptococcus intermedius	239	239	93%	3,0E-65	31,76%	677	WP_117286467.1
161 hypothetical protein [Bacteroidales bacterium]	Bacteroidales bacterium	292	292	90%	5,0E-85	31,69%	659	MBQ3245753.1
162 hypothetical protein [Verrucomicrobia bacterium]	Verrucomicrobia bacterium	298	298	97%	2,0E-87	31,66%	664	MCW5553404.1
163 hypothetical protein [Patescibacteria group bacterium]	Patescibacteria group bacterium	293	293	97%	2,0E-85	31,52%	664	MBU6214496.1
164 hypothetical protein A2W57_00770 [Candidatus Giovannonibacteria ba]	Candidatus Giovannonibacteria	286	286	92%	9,0E-83	31,45%	641	OGF74273.1
165 hypothetical protein [bacterium AH-315-M05]	bacterium AH-315-M05	286	286	90%	7,0E-83	31,43%	660	MBN4051515.1
166 hypothetical protein [Bacteroides sp. ET71]	Bacteroides sp. ET71	271	271	89%	2,0E-77	31,39%	645	MCL1614890.1
167 hypothetical protein [Sodaliophilus sp.]	Sodaliophilus sp.	281	281	89%	3,0E-81	31,38%	645	MDY5361326.1
168 hypothetical protein [Clostridium tyrobutyricum]	Clostridium tyrobutyricum	281	281	95%	9,0E-81	31,29%	671	WP_195509879.1
169 hypothetical protein [Spirochaetaceae bacterium]	Spirochaetaceae bacterium	298	298	98%	3,0E-87	31,27%	661	MDR1232002.1
170 hypothetical protein [Candidatus Saccharibacteria bacterium]	Candidatus Saccharibacteria bai	276	276	89%	6,0E-79	31,12%	671	MCA9346779.1
171 hypothetical protein DWW88_08360 [Bacteroides cellulosilyticus]	Bacteroides cellulosilyticus	279	279	89%	3,0E-80	31,09%	646	RGU27907.1
172 hypothetical protein HMPREF2531_02348 [Bacteroides intestinalis]	Bacteroides intestinalis	279	279	89%	5,0E-80	31,09%	670	XTS0244.1
173 TPA: hypothetical protein [Candidatus Saccharimonadales bacterium]	Candidatus Saccharimonadales	264	264	85%	2,0E-74	31,04%	671	HEX8763277.1
174 hypothetical protein [Candidatus Kerfeldbacteria bacterium]	Candidatus Kerfeldbacteria bacte	278	278	93%	1,0E-79	30,95%	670	MBI2985083.1
175 hypothetical protein [Clostridium sp.]	Clostridium sp.	292	292	95%	6,0E-85	30,93%	671	WP_333887885.1
176 hypothetical protein [Prevotella sp.]	Prevotella sp.	274	274	90%	3,0E-78	30,93%	650	MBQ4623394.1
177 hypothetical protein [Bacteroides fragilis]	Bacteroides fragilis	275	275	89%	8,0E-79	30,92%	646	WP_032591338.1
178 hypothetical protein A3D71_01720 [Candidatus Kaiserbacteria bacterium]	Candidatus Kaiserbacteria bacte	303	303	97%	4,0E-89	30,91%	665	OGG65371.1
179 hypothetical protein [Prevotella sp.]	Prevotella sp.	280	280	89%	1,0E-80	30,82%	645	MDY5091749.1
180 hypothetical protein [Rossellomorea marisflavi]	Rossellomorea marisflavi	260	260	93%	6,0E-73	30,76%	676	WP_318341953.1
181 Mva1269f restriction endonuclease [Kocuria varians]	Kocuria varians	266	266	98%	3,0E-75	30,20%	685	AAV97906.1
182 hypothetical protein [Verrucomicrobiota bacterium]	Verrucomicrobiota bacterium	299	299	95%	1,0E-87	29,17%	668	MCX6902957.1
183 hypothetical protein [Promineifilum sp.]	Promineifilum sp.	270	270	94%	2,0E-76	28,94%	680	MBX7251069.1
184 hypothetical protein [Amylibacter sp.]	Amylibacter sp.	189	189	85%	4,0E-47	28,71%	634	MDA9894940.1

Supplementary Table S2. List of the plasmids used for BsmI mutants expression. The pIVEX-CDF backbone is derived from the commercial high-copy pIVEX (Roche), the pUC origin of replication was changed for the p15A-compatible CloDF13 origin.

Plasmid name	ORF	Antibiotic resistance	Promoter	ori
pSB3K3-BsmIM	BsmIMA-PRP-BsmIMB	Kanamycin	pLac	p15A
pIVEX-CDF-BsmI	BsmI	Ampicilin	T7	CloDF13
pIVEX-CDF-Nb.BsmI	Nb.BsmI	Ampicilin	T7	CloDF13
pIVEX-N0.BsmI	N0.BsmI	Ampicilin	T7	pUC
pIVEX-Nt.BsmI	Nt.BsmI	Ampicilin	T7	pUC

Supplementary Table S3. List of the buffers used for the purification of BsmI and its mutants. Buffers A and B were used to perform an imidazole gradient during the elution of the affinity chromatography (HisTrap). Buffer C was used to dilute the obtained fractions to reduce the NaCl concentration down to 100 mM. Buffers D and E were used to perform a NaCl gradient during the elution of the ion exchange chromatography (Heparin) step presenting a gradient in the elution phase. F was the equilibration buffer of the size-exclusion chromatography (Superdex 200 Increase).

	A	B	C	D	E	F
KH ₂ PO ₄ -K ₂ HPO ₄ , pH 7.1	20 mM	20 mM	20 mM	20 mM	20 mM	-
Tris HCl, pH 7.5	-	-	-	-	-	10 mM
NaCl	500 mM	500 mM	-	100 mM	2 M	100mM
MgSO ₄	-	-	-	-	-	5 mM
EDTA, pH 8	0.1 mM	0.1 mM	0.1 mM	0.1 mM	0.1 mM	-
β-mercaptoethanol	10 mM	10 mM	10 mM	10 mM	10 mM	-
Glycérol	5 %	5 %	5 %	5 %	5 %	-
Imidazole	30 mM	500 mM	-	-	-	-

ADDITIONAL REFERENCES

1. Armalyte,E., Bujnicki,J.M., Giedriene,J., Gasiunas,G., Kosiński,J. and Lubys,A. (2005) Mva1269I: A Monomeric Type IIS Restriction Endonuclease from *Micrococcus Varians* with Two EcoRI- and FokI-like Catalytic Domains. *J. Biol. Chem.*, **280**, 41584–41594.
2. Goddard,T.D., Huang,C.C., Meng,E.C., Pettersen,E.F., Couch,G.S., Morris,J.H. and Ferrin,T.E. (2018) UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci.*, **27**, 14–25.
3. Crooks,G.E., Hon,G., Chandonia,J.-M. and Brenner,S.E. (2004) WebLogo: A Sequence Logo Generator. *Genome Res.*, **14**, 1188–1190.
4. Luscombe,N.M., Laskowski,R.A. and Thornton,J.M. (1997) NUCPLOT: A Program to Generate Schematic Diagrams of Protein-Nucleic Acid Interactions. *Nucleic Acids Res.*, **25**, 4940–4945.
5. Abramson,J., Adler,J., Dunger,J., Evans,R., Green,T., Pritzel,A., Ronneberger,O., Willmore,L., Ballard,A.J., Bambrick,J., *et al.* (2024) Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature*, **630**, 493–500.
6. Robert,X. and Gouet,P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.*, **42**, W320–W324.