The CRISPR/Cas bacterial immune system cleaves phage and plasmid DNA

Josiane E. Garneau[1], Marie-Ève Dupuis[1], Manuela Villion[1], Dennis A. Romero[2], Rodolphe Barrangou[2], Patrick Boyaval[3], Christophe Fremaux[3], Philippe Horvath[3], Alfonso H. Magadán[1] and Sylvain Moineau[1]*

[1]*Département de biochimie, de microbiologie et de bio-informatique,*

Faculté des sciences et de génie,

Groupe de recherche en écologie buccale, Faculté de médecine dentaire,

Félix d'Hérelle Reference Center for Bacterial Viruses,

Université Laval, Quebec City, Quebec, Canada, G1V 0A6.

[2] *Danisco USA Inc., 3329 Agriculture Drive, Madison, WI 53716, USA.*

[3] *Danisco France SAS, Boîte Postale 10, F-86220 Dangé-Saint-Romain, France.*

E-mail addresses: josiane.garneau.1@ulaval.ca; marie-eve.dupuis.2@ulaval.ca; manuela.villion.1@ulaval.ca; dennis.romero@danisco.com; rodolphe.barrangou@danisco.com; patrick.boyaval@danisco.com; christophe.fremaux@danisco.com; philippe.horvath@danisco.com; ahmagadan@gmail.com; sylvain.moineau@bcm.ulaval.ca

* Corresponding author contact information: [1]Département de biochimie et de microbiologie, Faculté des sciences et de génie, Université Laval, Quebec City, Quebec, Canada, G1V 0A6

Phone: 418-656-3712. Fax: 418-656-2861. E-mail: Sylvain.Moineau@bcm.ulaval.ca

Bacteria and Archaea have developed multiple defence strategies against foreign nucleic acids such as viral genomes and plasmids. Among them, clustered regularly interspaced short palindromic repeats (CRISPR) loci, together with *cas* (CRISPR associated) genes, form the CRISPR/Cas immune system, which involves partially palindromic repeats separated by short stretches of DNA called spacers, acquired from extrachromosomal elements. It was recently demonstrated that these variable loci can incorporate spacers from infecting phages and then provide immunity against subsequent phage infections in a sequence-specific manner. Here, we show that the *Streptococcus thermophilus* CRISPR1/Cas system can also naturally acquire spacers from a self-replicating plasmid containing an antibiotic resistance gene, leading to plasmid loss. Acquired spacers that match antibiotic resistance genes provide a novel means to naturally select bacteria that cannot uptake and disseminate such genes. We also provide the first *in vivo* evidence that the CRISPR1/Cas system specifically cleaves plasmid and phage double-stranded DNA within the proto-spacer, at specific sites. Altogether, the CRISPR/Cas immune system is remarkably adapted to rapidly cleave invading DNA and could be exploited to generate safer microbial strains.

Clustered regularly interspaced short palindromic repeats (CRISPR) loci were discovered in 1987 in *Escherichia coli.* However, the interest in these genetic elements increased in the early 2000s, as they were identified, along with many CRISPR-associated (Cas) proteins, in several Prokaryotes[1-3]. Recently, it was shown that the short spacers (21-72 bp) between these repeats originated from extrachromosomal DNA[4-7]. Most importantly, it was experimentally demonstrated that those short spacers can provide resistance against phage infection and plasmid transformation[8-10].

The CRISPR/Cas immune systems act in two steps: i) the adaptation stage, where new spacers derived from foreign DNA (proto-spacers) are generally acquired at the leader-end of the CRISPR locus[11,12], and ii) the interference stage, where the CRISPR/Cas system targets either invading DNA[10] or RNA[13]. The mechanistic details of spacer acquisition are still unknown, but a clearer picture is emerging for the interference stage, which starts with the transcription of the CRISPR locus from a promoter located within the leader sequence[14,15]. The full length RNA is subsequently cleaved by a protein or protein complex, generating short CRISPR RNAs (crRNAs)[16-19]. In *Pyrococcus*, Cas proteins use the crRNAs to target foreign RNA by complementarity in a ruler-anchored manner[13]. However, the *in vivo* mechanism of plasmid and viral interference has yet to be determined.

We previously showed that when phage-sensitive *Streptococcus thermophilus* cells are infected by virulent phages, a sub-set of cells (frequency of $< 10^{-6}$) naturally diversify into bacteriophage-insensitive mutants (BIMs) through the acquisition of novel spacers derived from the invading phage genome into CRISPR1 and/or CRISPR3 (ref. [9,11,20]). Here, we investigate the *in vivo* activity of the CRISPR/Cas system in *S. thermophilus* against both phage and plasmid DNA.

CRISPR/Cas affects plasmid stability

In *silico* analyses previously suggested that spacers may also be derived from *S. thermophilus* plasmids[11]. Because the CRISPR/Cas system was recently demonstrated to interfere with plasmid transfer in *Staphylococcus*[10], we put forward the hypothesis that the CRISPR/Cas system is responsible for the scarcity of plasmids in wild-type strains of *S. thermophilus*[21]. To test this hypothesis, we first introduced by electroporation the vector pNT1 (ref. [22]) into the plasmid-free *S. thermophilus* strain DGCC7710. This vector, which is derived from a native *S. thermophilus* plasmid, replicates via a rolling-circle mode and carries an added chloramphenicol resistance (*cat*) gene as selection marker. A representative transformant was grown in liquid medium for about 60 generations in the absence of chloramphenicol and aliquots were screened for antibiotic-sensitive colonies.

A total of 54 colonies (out of 900 tested) became sensitive to chloramphenicol and concomitantly lost pNT1. Thirty of these 54 colonies acquired a new spacer-repeat unit in CRISPR1. No spacer was integrated in the three other CRISPR loci of *S. thermophilus* DGCC7710 (ref.[8]). Sequence analysis of the CRISPR1 in the aforementioned 30 colonies identified 14 different spacers (S43-S56), all of which were homologous to pNT1 sequences (Table 1, Fig. 1). Plasmid stability assays were also performed using two isogenic DGCC7710 strains in which *cas5* (*csn1*-like) or *cas7* genes associated to CRISPR1 were inactivated prior to the introduction of pNT1. Plasmid pNT1 was highly stable in the DGCC7710::p*cas5*$^-$ mutant as no chloramphenicol-sensitive colonies could be isolated, out of 1800 screened. Of 170 randomly selected chloramphenicol-resistant colonies, none had acquired a new spacer in CRISPR1. On the other hand, chloramphenicol-sensitive colonies were readily obtained with the strain DGCCC7710::p*cas7*$^-$, but none of the 200 colonies tested had acquired a new spacer in CRISPR1, suggesting that plasmid loss was likely the result of

other mechanisms responsible for plasmid instability[23,24]. These data are consistent with previous findings that *cas5* is involved in the interference stage[9] and its inactivation favors plasmid stability while *cas7* is linked to spacer acquisition[9].

Taken altogether, these observations indicate that the CRISPR/Cas system causes plasmid loss in *S. thermophilus*.

CRISPR/Cas targets antibiotic resistance genes

To assess whether the acquired spacers from pNT1 cause plasmid interference in these plasmid interfering mutants (PIMs), we tested their propensity for pNT1 re-introduction. We were unable to obtain transformants of the PIMs that carried a spacer perfectly matching a proto-spacer associated with the consensus proto-spacer adjacent motif (PAM) NNAGAAW (Table 1). We previously showed that the newly added spacer must be identical to the proto-spacer in the phage genome to confer phage resistance, and that a PAM located downstream of the proto-spacer is required for the resistance phenotype[9,20,25,26]. Here, approximately half of the proto-spacer adjacent motif sequence contained one or two nucleotide mismatches. This tolerance for PAM degeneracy could be due to the lower selective pressure for plasmids as compared to phages.

Some plasmid interfering mutants targeting proto-spacers associated with non-consensus motifs (**N**NNAGAA**G**, NNA**T**AAA, NN**G**GAAT or NNAGAA**G)** were also refractory to pNT1 reintroduction. Interestingly, a plasmid carrying the degenerate motif NNA**T**AAA downstream of proto-spacer S52 could not be transformed into the corresponding plasmid interfering mutants, whereas phages carrying the same PAM could infect the matching BIMs[20] (Sup. Table 1). It is also worth mentioning that one plasmid interfering mutant contained a spacer (S47) which matched the last 29 nucleotides (out of 30) of the corresponding proto-spacer in pNT1, suggesting that

sequence identity at the 5'-end of the spacer might be less important than in the middle or at the 3'-end.

Two plasmid interfering mutants (DGCC7710$_{pNT1}^{+S55}$ and DGCC7710$_{pNT1}^{+S56}$, also named PIM S55 and PIM S56 for simplicity) carried a spacer targeting the *cat* gene (S55) or its downstream region (S56) (Fig. 1, Table 1). These spacers rendered the cells untransformable with pNT1 and also with another plasmid (pTRK687) carrying the same *cat* gene. Thus, we conclude that the CRISPR/Cas system provides a simple and natural means to develop a bacterial strain, which is refractory to the acquisition of plasmids that carry antibiotic resistant genes.

Degeneracy of the motif influences plasmid interference

Two other plasmid interfering mutants (PIM S45 and PIM S46) carrying a spacer associated with a non-consensus proto-spacer adjacent motif could be re-transformed with pNT1 but at lower frequencies (Table 1). Unexpectedly, the re-transformed PIM S46 visibly carried a linear form of pNT1 (Fig. 2A). *S. thermophilus* strain PIM S46 also contained the usual circular form of the vector, albeit at much lower concentration than the linear form (Fig. 2B), and could still grow in a medium containing chloramphenicol. However, 74-100% of the cells lost pNT1 within 14 generations after growth in a non-selective medium. Under the same conditions, 0-10% of the wild-type colonies lost the plasmid. Therefore, although the NNAG**T**A**G** motif was initially permissive for pNT1, the CRISPR/Cas machinery still eliminated the circular and linear plasmid forms within a few generations. The reason for the relative stability of the linear form of the plasmid, under selective pressure, is still unclear.

To assess whether the observed plasmid linearization was the result of the CRISPR/Cas activity, the *cas5* and *cas7* genes of PIM S46 were also inactivated and the isogenic strains transformed with pNT1. Plasmid content analysis revealed the presence

of only the circular form in PIM S46::p*cas5*$^{-}$, confirming the involvement of *cas5* in plasmid interference. On the other hand, pNT1 was still linear in mutant PIM S46::p*cas7*$^{-}$ and in PIM S46 (Fig 2A).

Then, we investigated the terminal ends of the pNT1 DNA molecules by directly sequencing the linear plasmid extracted from the PIM S46 strain. We unambiguously determined that pNT1 had been cut once (blunt), within the S46 proto-spacer, after the 27th nucleotide, 3 bases upstream of the proto-spacer adjacent motif (Table 2) (Fig. 2C). We conclude that double-stranded plasmid DNA is targeted and cleaved within the proto-spacer by the CRISPR/Cas machinery.

The CRISPR/Cas system cleaves phage DNA in the proto-spacers

To test whether the CRISPR/Cas system cleaves other invading DNA within the proto-spacer, we analyzed the fate of phage DNA during the infection of bacteriophage-insensitive mutants (BIMs) containing phage-derived spacers (Fig. 3A). We used the virulent streptococcal phage 2972 (ref.[27]), its host DGCC7710, and three previously described BIMs derived from DGCC7710 following the challenge with virulent phages 2972 and/or 858, namely DGCC7710$_{\Phi 2972}^{+S4}$, DGCC7710$_{\Phi 2972}^{+S7}$, and DGCC7710$_{\Phi 2972}^{+S4}{}_{\Phi 858}^{+S32}$ (BIM S4, BIM S7, and BIM S4/S32, respectively)[20]. These bacterial strains were infected with phage 2972 and total DNA was extracted from phage-infected cells sampled overtime. Southern hybridizations of the DNA from the time-course infection of the BIMs with probes targeting regions close to the proto-spacers revealed that the phage DNA is rapidly cleaved by the CRISPR/Cas machinery (Fig. 3B, C). The cleavage site in the phage genome appeared to be within or in the vicinity of the proto-spacer during the infection of BIMs, while no cleavage of phage DNA was observed in the sensitive strain DGCC7710 (Fig. 3C). The cleavage patterns were similar in the phage-infected BIMs, even though the S4 and S7 spacers are derived

from different DNA strands and transcriptional modules of the phage genome[28] (Fig. 3A). We also conducted the same experiments with isogenic derivatives of *S. thermophilus* BIM S4, in which *cas5* and *cas7* genes were independently inactivated. The inactivation of the *cas5* gene restored phage sensitivity (efficiency of plaquing of 1) and no phage DNA cleavage occurred, confirming the involvement of *cas5* in interference. In contrast, the inactivation of *cas7* did not affect phage resistance (efficiency of plaquing of 10^{-6}) and phage DNA was cleaved (Fig. 3C).

We also investigated the fate of phage DNA during the infection of *S. thermophilus* BIM S4/S32, which contains two new spacers that target phage 2972. The phage genome was cleaved within each proto-spacer in this infected bacteriophage-insensitive mutant (Fig. 3C). The 5.2-kb and 5.8-kb bands corresponded to a cleavage at the S4 or S32 proto-spacers, respectively, whereas the 1.4-kb band resulted from cleavage at both sites. This is consistent with the previous observation that resistance increases with the number of spacers acquired[20]. Altogether, our results show that the *in vivo* target of the *S. thermophilus* CRISPR1/Cas system is DNA for both plasmid (Fig. 2) and phage (Fig. 3).

We then wanted to determine whether the phage DNA cleavage site was identical to that of the plasmid. Contrary to the method used for the determination of the cleavage site of plasmid pNT1, the low amount of cleaved phage DNA in the infected bacteriophage-insensitive mutants rendered direct sequencing of the extremities impossible. An alternative sequencing method was used (see Sup. Methods and Sup. Fig. 1) as the cleaved 1.4-kb fragment (from *S. thermophilus* BIM S4/S32) was subjected to inverse PCR to obtain a product which joins the ends of the S4- and S32-cleaved proto-spacers. Sequence analyses of the amplicon revealed that the phage fragment had indeed been cleaved within S4 and S32. For the other cleavage sites, the extracted phage DNA was digested *in vitro* by restriction endonucleases producing

blunt-end and cutting upstream and downstream of each proto-spacer (see Sup Fig. 1 and 2). Fragments comprising the cleaved proto-spacer at one extremity and a blunt site at the other were ligated and inverse PCRs were performed as described previously. PCR products were obtained for every fragment and sequenced.

Analysis of the amplicons revealed that all the cleavage sites were within the proto-spacers (S4, S7, and S32). In every cleaved proto-spacer analyzed, one cleavage site was located after the 27th nucleotide, 3 bases upstream of the proto-spacer adjacent motif, as observed for the linearized plasmid in PIM S46 (Table 2). Curiously, a second cleavage site was detected within the proto-spacers but only for the spacers targeting the positive strand of the phage genome (see S32 and S4 in Table 2). This second site was located 19 or 20 nucleotides upstream of the PAM. Of interest, the plasmid proto-spacer S46 described above was on the negative strand and was cut only once (Table 2). To confirm these results, three additional phage-infected bacteriophage-insensitive mutants that have acquired different spacers (S40, S41, or S42) were analyzed similarly. The single (negative strand) and double (positive strand) cleavage sites within the proto-spacer were confirmed in these three bacteriophage-insensitive mutants (Table 2). Of note, the two cleavage sites within the S40 proto-spacer (a natural 5'-end truncated version of the S4 proto-spacer with 29 nucleotides instead of 30) were at the same position, suggesting that the CRISPR1/Cas system in *S. thermophilus* likely acts in a 3'-end ruler-anchored manner. Ligation and amplification of each cleaved fragment could be obtained and no nucleotide was missing in the S7, S41, S42, and S46 proto-spacer sequences, confirming the blunt-end cleavage activity of *S. thermophilus* CRISPR1/Cas system.

Overall, we have established that the *S. thermophilus* CRISPR1/Cas system cleaves both phage and plasmid DNA *in vivo*. This endonuclease activity, which appears to require *cas5,* is proto-spacer-specific and orientation dependent. In that

regard, the *S. thermophilus* CRISPR1/Cas interference mechanism differs from that recently described in *Pyrococcus furiosus*, whereby RNA is the target[13] but corroborates the previous work on *Staphylococcus* which identified DNA as the target[10]. Furthermore, we show that the DNA cleavage activity is responsible for both phage resistance in BIMs and plasmid instability in PIMs. The antagonism between plasmid maintenance and CRISPR/Cas activity likely explains the natural scarcity of plasmids in *S. thermophilus*. The CRISPR/Cas immune system can be leveraged to naturally generate safer and more robust organisms with increased phage resistance and able to interfere with the dissemination of plasmids that carry antibiotic resistance markers.

Methods summary

Microbial conditions. *S. thermophilus* strains were grown at 37°C or 42°C in LM17. Strains were infected with phage 2972 at a multiplicity of infection of 5 (ref. [20]). *cas5* and *cas7* genes were inactivated as described[9]. The identity of *S. thermophilus* strains was confirmed by sequencing the CRISPR loci[8] and by phage assays[20].

Molecular biology. DNA was extracted from bacterial cells as described[29]. DNA probes for Southern[30] were prepared with the PCR Dig labeling mix (Roche) (Sup. Table 2). Hybridization and detection (CDP-star) were performed as recommended by Roche.

Plasmid stability. *S. thermophilus* DGCC7710 (or its *cas5*$^-$ and *cas7*$^-$ derivatives) transformed with pNT1 (ref. [22]) was used to inoculate 10 mL of LM17. One hundred microliters of the previous culture was inoculated into 10 mL of fresh LM17 medium every morning (42˚C) and night (37˚C) for 5 days, for a total of 9 inoculations. For each culture, 100 colonies were screened for chloramphenicol sensitivity. The CRISPR1 of the chloramphenicol-sensitive clones was analyzed. PIMs with new spacers were electroporated[31] with ~600 ng of pNT1 or pTRK687 (ref. [32]). PIMs S55 and S56 were also transformed with pLS1 (ref. [33]). Plasmid content of re-transformed PIMs S45 and S46 was verified by extraction (Qiagen) and extremities of the linearized plasmid from PIM S46 were sequenced.

Cleavage site of phage DNA. Total DNA from phage-infected BIMs was extracted[29] and digested using endonucleases generating blunt ends (Sup. Fig. 2). After migration, agarose bands corresponding to both sides of the cleaved proto-spacers were recovered, DNA eluted and used for intramolecular ligation O/N at 16°C, with T4 DNA ligase

(Invitrogen). Ligation product was used as a template for inverse PCR[34] (Sup. Fig. 1 and 2, Sup. Table 3) and amplicons were sequenced (primers in Sup. Table 3).

Full Methods and any associated references are available in the online version of the paper at **www.nature.com/nature**.

REFERENCES

1. Ishino, Y., Shinagawa, H., Makino, K., Amemura, M. & Nakata, A. Nucleotide sequence of the *iap* gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *J. Bacteriol.* **169,** 5429-5433 (1987).
2. Jansen, R., van Embden, J. D. A., Gaastra, W. & Schouls, L. M. Identification of genes that are associated with DNA repeats in prokaryotes. *Mol. Microbiol.* **43,** 1565-1575 (2002).
3. Mojica, F. J., Díez-Villaseñor, C., Soria, E. & Juez, G. Biological significance of a family of regularly spaced repeats in the genomes of Archaea, Bacteria and mitochondria. *Mol. Microbiol.* **36,** 244-246 (2000).
4. Bolotin, A., Quinquis, B., Sorokin, A. & Ehrlich, S. D. Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology* **151,** 2551-2561 (2005).
5. Mojica, F. J., Díez-Villaseñor, C., García-Martínez, J. & Soria, E. Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J. Mol. Evol.* **60,** 174-182 (2005).

6. Pourcel, C., Salvignol, G. & Vergnaud, G. CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology* **151,** 653-663 (2005).

7. Andersson, A. F. & Banfield, J. F. Virus population dynamics and acquired virus resistance in natural microbial communities. *Science* **320,** 1047-1050 (2008).

8. Horvath, P. & Barrangou, R. CRISPR/Cas, the immune system of Bacteria and Archaea. *Science* **327,** 167-170 (2010).

9. Barrangou, R. *et al.* CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315,** 1709-1712 (2007).

10. Marraffini, L. A. & Sontheimer, E. J. CRISPR interference limits horizontal gene transfer in staphylococci by targeting DNA. *Science* **322,** 1843-1845 (2008).

11. Horvath, P. *et al.* Diversity, activity, and evolution of CRISPR loci in *Streptococcus thermophilus*. *J. Bacteriol.* **190,** 1401-1412 (2008).

12. Marraffini, L. A. & Sontheimer, E. J. CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nature Rev. Genet.* **11,** 181-190 (2010).

13. Hale, C. R. *et al.* RNA-guided RNA cleavage by a CRISPR RNA-Cas protein complex. *Cell* **139,** 945-956 (2009).

14. Lillestøl, R. K. *et al.* CRISPR families of the crenarchaeal genus *Sulfolobus*: bidirectional transcription and dynamic properties. *Mol. Microbiol.* **72,** 259-272 (2009).

15. Pul, U. *et al.* Identification and characterization of *E. coli* CRISPR-*cas* promoters and their silencing by H-NS. *Mol. Microbiol.* **75,** 1495-1512 (2010).

16. Tang, T. H. *et al.* Identification of 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus*. *Proc. Natl Acad. Sci. USA* **99,** 7536-7541 (2002).

17. Carte, J., Wang, R. Y., Li, H., Terns, R. M. & Terns, M. P. Cas6 is an endoribonuclease that generates guide RNAs for invader defense in prokaryotes. *Genes Dev.* **22,** 3489-3496 (2008).

18. Hale, C., Kleppe, K., Terns, R. M. & Terns, M. P. Prokaryotic silencing (psi)RNAs in *Pyrococcus furiosus*. *RNA* **14,** 2572-2579 (2008).

19. Brouns, S. J. J. *et al.* Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321,** 960-964 (2008).

20. Deveau, H. *et al.* Phage response to CRISPR-encoded resistance in *Streptococcus thermophilus*. *J. Bacteriol.* **190,** 1390-1400 (2008).

21. Turgeon, N. & Moineau, S. Isolation and characterization of a *Streptococcus thermophilus* plasmid closely related to the pMV158 family. *Plasmid* **45,** 171-183 (2001).

22. Vaillancourt, K. *et al.* Role of *galK* and *galM* in galactose metabolism by *Streptococcus thermophilus*. *Appl. Environ. Microbiol.* **74,** 1264-1267 (2008).

23. Girard, S. L. & Moineau, S. Analysis of two theta-replicating plasmids of *Streptococcus thermophilus*. *Plasmid* **58,** 174-181 (2007).

24. Kiewiet, R., Kok, J., Seegers, J. F., Venema, G. & Bron, S. The mode of replication is a major factor in segregational plasmid instability in *Lactococcus lactis*. *Appl. Environ. Microbiol.* **59,** 358-364 (1993).

25. Bolotin, A. *et al.* Complete sequence and comparative genome analysis of the dairy bacterium *Streptococcus thermophilus*. *Nature Biotechnol.* **22,** 1554-1558 (2004).

26. Mojica, F. J., Díez-Villaseñor, C., García-Martínez, J. & Almendros, C. Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology* **155,** 733-740 (2009).

27. Lévesque, C. *et al.* Genomic organization and molecular analysis of virulent bacteriophage 2972 infecting an exopolysaccharide-producing *Streptococcus thermophilus* strain. *Appl. Environ. Microbiol.* **71,** 4057-4068 (2005).

28. Duplessis, M., Russell, W. M., Romero, D. A. & Moineau, S. Global gene expression analysis of two *Streptococcus thermophilus* bacteriophages using DNA microarray. *Virology* **340,** 192-208 (2005).

29. Fortier, L.-C. & Moineau, S. Morphological and genetic diversity of temperate phages in *Clostridium difficile*. *Appl. Environ. Microbiol.* **73,** 7358-7366 (2007).

30. Sambrook, J. & Russell, D. W. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 2001).

31. Buckley, N. D., Vadeboncoeur, C., LeBlanc, D. J., Lee, L. N. & Frenette, M. An effective strategy, applicable to *Streptococcus salivarius* and related bacteria, to enhance or confer electroporation competence. *Appl. Environ. Microbiol.* **65,** 3800-3804 (1999).

32. Sturino, J. M. & Klaenhammer, T. R. Expression of antisense RNA targeted against *Streptococcus thermophilus* bacteriophages. *Appl. Environ. Microbiol.* **68,** 588-596 (2002).

33. Nieto, C., Fernández de Palencia, P., López, P. & Espinosa, M. Construction of a tightly regulated plasmid vector for *Streptococcus pneumoniae*: controlled expression of the green fluorescent protein. *Plasmid* **43,** 205-213 (2000).

34. Ochman, H., Gerber, A. S. & Hartl, D. L. Genetic applications of an inverse polymerase chain reaction. *Genetics* **120,** 621-623 (1988).

35. Clark, J. M. Novel non-templated nucleotide addition reactions catalyzed by procaryotic and eucaryotic DNA polymerases. *Nucleic Acids Res.* **16,** 9677-9686 (1988).

Full methods

Bacterial, phage, and culture conditions. *Streptococcus thermophilus* strain DGCC7710 (ref. [9]) and BIMs or PIMs were grown in M17 broth supplemented with 0.5% lactose (LM17) at 37˚C or 42°C. Phages were propagated in LM17 supplemented with 10 mM $CaCl_2$. Purified phage preparations (10^{11}-10^{12} PFU/mL) were obtained by ultracentrifugation using a discontinuous CsCl gradient[30]. Phage sensitivity of the isolated BIMs was estimated as reported previously[36]. *Cas5* and *cas7* genes were inactivated as described previously[9] in three *S. thermophilus* strains : DGCC7710, BIM S4, and PIM S46. All *cas5*$^-$ and *cas7*$^-$ derivatives were grown under erythromycin selective pressure. Efficiency of plaquing was determined as reported[20].

Phage infection. Bacterial strains were incubated at 42°C in 100 mL of LM17 to an optical density of 0.5 at 600 nm. Cultures were concentrated by centrifugation and pellets were resuspended in 10 mL of preincubated LM17 medium containing 10 mM $CaCl_2$. After removing a 1-mL uninfected sample, each bacterial culture was infected with purified phage 2972 at a multiplicity of infection (MOI) of 5, and incubated at 42°C. Samples were taken after 2, 15, 30, and 45 minutes, centrifuged for 15 seconds at 16,000*g* and pellets were flash frozen, and stored at -80°C until DNA extraction. The latent period of phage 2972 was previously determined to be 34 min on *S. thermophilus* DGCC7710. The maximum burst of phages is 40 min after the beginning of the infection[28].

DNA extraction and Southern hybridization. Total DNA extractions were performed as described elsewhere[29], with these modifications: 25% (w/v) sucrose and 60 mg/mL

lysozyme were used as well as only one phenol-chloroform extraction. Dissolved DNA solutions were treated with 1 µg of RNase, incubated 15 minutes at 37°C, and their concentration determined using the Nanodrop 2000 spectrophotometer (Thermo Scientific).

Ten µg of BIM DNA and 1 µg of wild-type DNA were digested with either HpaII or BclI (Roche). As controls, 10 µg of bacterial DNA and 10 ng of phage DNA were digested and were used for probe hybridizations. In addition, 500 ng of phage DNA was used as a digestion and transfer control. DNA fragments were separated in a 0.8% agarose gel in 1X TAE buffer, stained with EZ-Vision DNA dye (Amresco) and photographed under UV. DNA was transferred onto positively charged nylon membranes (Roche) by capillary blotting[30]. Membranes were stained for 45 seconds with methylene blue, rinsed with distilled water, and dried in ambient air. DNA probes were prepared with PCR Dig labeling mix (Roche) (Sup. Table 2). Prehybridization, hybridization, washes, and detection (CDP-star) were performed as recommended by Roche.

Plasmid stability and PIM analysis. *S. thermophilus* DGCC7710 (or the *cas5*$^-$ and *cas7*$^-$ isogenic derivatives) transformed with pNT1 plasmid[22] (GenBank accession number HQ010044) was picked from a LM17 plate containing 5 µg/mL of chloramphenicol. Plasmid and CRISPR contents of this transformant were confirmed by sequencing and used to inoculate 10 mL of LM17. One hundred microliters of the previous culture was inoculated into 10 mL of fresh LM17 medium every morning and night for 5 days, for a total of 9 inoculations. Cultures were grown at 37ºC for overnight incubations and at 42ºC for 8 hours during the day. Finally, these cultures were serially diluted and plated on LM17. For each culture, 100 colonies were screened for the absence of the antibiotic resistance phenotype. Plasmid stability of pNT1 was analyzed

by comparing the number of colonies appearing on selective agar (5 μg/ml of chloramphenicol) versus the number of colonies on LM17 agar. The CRISPR1 of chloramphenicol-sensitive clones was verified: The 5'-end of the CRISPR1 of the PIMs or BIMs was amplified by PCR with the primers yc70 (5'-tgctgagacaacctagtctctc-3')[4] and RDS7revBamHI (5'-ggatccggatccgttgaggccttgttc-3'), and sequenced using the same primers. Fourteen different PIMs having acquired a new spacer from pNT1 were electroporated as described elsewhere[31] with ~600 ng of pNT1 or pTRK687 (ref. [32]) vectors. All PIMs (except for PIM S55 and PIM S56) could be transformed with the control vector pTRK687, which is carrying a chloramphenicol-resistance gene as selection marker. In the case of PIM S55 and PIM S56, they could be transformed with pLS1, which is carrying an erythromycin-resistance gene as a selection marker[33].

The proto-spacer region of pNT1 isolated from the pNT1-transformed PIM S45 was amplified and sequenced using the primers pNT1_104 (5'-gtgccttgaaccttagagccacaa-3') and NT17215_3 (5'-gttcagagtatggactgccg-3'). Chloramphenicol resistant clones were also checked for spacer acquisition. Their CRISPR1 was amplified by PCR using the primers yc70 and RDS7revBamHI cited above.

Plasmid extraction and plasmid cleavage site determination. The plasmid content of *S. thermophilus* strains was isolated using QIAquick Spin Miniprep columns (Qiagen). The QIAquick protocol was modified by treating the cells with P1 buffer containing sucrose (25%) and lysozyme (60 mg/mL), and incubating at 37°C for 15 to 30 min. Extremities of the linearized plasmid were directly sequenced with the primers NT17225_2 (5'-tacgtacttgtgttactattg-3') and NT17215_3 (5'-cggcagtccatactctgaac-3').

Phage cleavage site determination. Each bacteriophage-insenstive mutant (BIM) was independently infected for 45 min with phage 2972 at a multiplicity of infection of 5. Total DNA from phage-infected BIMs was extracted and digested using endonucleases cutting upstream and downstream each proto-spacer and generating blunt ends (Roche and NEB) (Sup. Fig. 2). DNA was migrated on a 0.8% agarose gel. Bands corresponding to the left and right fragments of the cleaved proto-spacer were extracted from the gel (band sizes estimated according to cleavage within the proto-spacer). DNA was extracted using QIAEX II Gel Extraction Kit (Qiagen). Each fragment comprised the cleaved proto-spacer at one extremity and a blunt site at the other. Total eluted product (50 µl) of each fragment was used for intramolecular ligation O/N at 16°C, with T4 DNA ligase (Invitrogen). In the case of the BIM S4/S32, the fragment delimited by the cleaved proto-spacers S4 and S32 could be directly used for intramolecular ligation. Five µL of the ligation product were used as a template for PCR (100 µL reaction volume), with Taq Polymerase (Roche) (Sup. Table 3). The cycling conditions were: 94ºC/45 sec, 55ºC/45 sec, 72ºC/1 min for 35 cycles with a final extension of 72ºC/10 min[34]. PCR products were sequenced using primers listed in Sup. Table 3.

Sequences. All sequencing reactions were performed by the "Plateforme de Séquençage et de Génotypage des Génomes" service from the CHUL-CHUQ Research Center. The sequences from both strands were analyzed using BioEdit software (http://www.mbio.ncsu.edu/BioEdit/bioedit.html) or Staden package (http://staden.sourceforge.net/staden_home.html).

SUPPLEMENTARY REFERENCES

36. Moineau, S. *et al.* Characterization of lactococcal bacteriophages from Québec cheese plants. *Can. J. Microbiol.* **38,** 875-882 (1992).

Supplementary information accompanies the paper on **www.nature.com/nature.**

Acknowledgements. We thank T.R. Klaenhammer for pTRK687, Barb Conway for editorial assistance, Denise Tremblay, Audrey Fleury, and Buffy Stahl for technical assistance. M.-E.D. is the recipient of a Novalait/FQRNT graduate scholarship and A.H.M. of a Clarín/FICYT postdoctoral scholarship. S.M. acknowledges funding from NSERC (Discovery program) and CIHR (Team Grant - Emerging: Novel Alternatives to Antibiotics).

Author Contributions. S.M. conceived and headed the project. All the authors participated in the design of the study. J.G., M.E.D., M.V., and A.M. performed the experiments. J.G. and S.M wrote the main parts of the manuscript. All the authors analyzed the results and commented on the manuscript.

Author Information. Reprints and permissions information is available at npg.nature.com/reprints. D.A.R, R.B., P.B., C.F., and P.H. have submitted patent applications relating to various uses of CRISPR. Correspondence and requests for materials should be addressed to S.M. (Sylvain.Moineau@bcm.ulaval.ca).

TABLES

Table 1. PIMs obtained after growth of *S. thermophilus* DGCC7710 containing pNT1 in a medium without antibiotic.

PIM	Spacer acquired[a] (5' - 3')	PAM[b]	Gene[c]	Position on pNT1	Strand	Transformation with pNT1[d] (CFU/µg DNA)	Number of times isolated[e]
DGCC7710$_{pNT1}^{+S43}$	GATCAAATAACTAATAAATACCCAGTACTT	**TGC**AGAA**G**	-	29	-	-	1
DGCC7710$_{pNT1}^{+S44}$	GACCCCCCTTTTAAGTGCCGAGTGCCAAAT	TGAGAAA	*dso*	63	+	-	1
DGCC7710$_{pNT1}^{+S45}$	TATACTTGGGTTAATTATACCGTATGGCAA	**A**AGAAA	-	436	+	4.4 ± 3.5	1
DGCC7710$_{pNT1}^{+S46}$	TTTCCCAATCTTCTGGAATTGAATCGGGAT	AGAG**T**A**G**	*rep*	528	-	3.7 ± 2.6 X 10^2	2
DGCC7710$_{pNT1}^{+S47}$	**C**ATGATCTGCAATAATATTGCAGACCTCGT	CTAGAAT	*rep*	917	-	-	1
DGCC7710$_{pNT1}^{+S48}$	GATGATCTGCAATAATATTGCAGACCTCGT	CTAGAAT	*rep*	918	-	-	2
DGCC7710$_{pNT1}^{+S49}$	CGATGATCTGCAATAATATTGCAGACCTCGT	CTAGAAT	*rep*	919	-	-	1
DGCC7710$_{pNT1}^{+S50}$	AATTTAGTTCCGTCAGTAGATTATGAAACT	GGAGAA**G**	*rep*	1074	+	-	1
DGCC7710$_{pNT1}^{+S51}$	AAAAGCAATGAGTTACATGGTTGCAAGAAT	GCAGAAA	*mob*	1488	+	-	4
DGCC7710$_{pNT1}^{+S52}$	GCCCCAGCTTACTATCAAGGAGCTTTCACG	GCA**T**AAA	*sso*	1999	-	-	1
DGCC7710$_{pNT1}^{+S53}$	CGCCACAGGTTACTTGCTGTCAAGGAGACC	AT**G**GAAT	-	2066	+	-	6
DGCC7710$_{pNT1}^{+S54}$	TCGTTTGTTGAACTAATGGGTGCTTTAGTT	GAAGAAT	-	2246	-	-	4
DGCC7710$_{pNT1}^{+S55}$	AGAGTTTTATGATTTATACCTTTCTGATGT	AGAGAAA	*cat*	2717	+	-	3
DGCC7710$_{pNT1}^{+S56}$	TTCTTCAACTAACGGGGCAGGTTAGTGACA	TTAGAAA	-	3114	-	-	2
DGCC7710	–					1.4 ± 0.6 X 10^3	

[a] Nucleotide mismatch with pNT1 sequence is bolded and underlined.

[b] Nucleotide mismatches with the CRISPR1 PAM (NNAGAAW) are bolded and underlined.

[c] – intergenic region.

[d] – indicates that pNT1 could not be electroporated into the PIMs (<1 CFU/µg DNA), n ≥ 2;

[e] Indicates the number of times each PIM was isolated in the assay.

Table 2. Cleavage sites within the proto-spacers in phage 2972 or plasmid pNT1.

BIM or PIM	Spacer (size)	Homology (position)	Strand	Proto-spacer (5' - 3')	PAM[a]
DGCC7710$_{\Phi2972}^{+S4}$	S4 (30 nt)	Phage 2972 (31582)	+	CTCAGTCGTT↓ACTGGTGAACCAGTTTC↓AAT	TGAGAAA
DGCC7710$_{\Phi2972}^{+S4}{}_{\Phi858}^{+S32}$	S4 (30 nt)	Phage 2972 (31582)	+	CTCAGTCGTT↓ACTGGTGAACCAGTTTC↓AAT	TGAGAAA
DGCC7710$_{\Phi2972}^{+S4}{}_{\Phi858}^{+S32}$	S32 (30 nt)	Phage 2972 (33044)	+	ATTGTCTATTA↓CGACAACATGGAAGAT↓GAT	GTAGAAA
DGCC7710$_{\Phi2972}^{+S40}$	S40 (29 nt)	Phage 2972 (31583)	+	TCAGTCGTT↓ACTGGTGAACCAGTTTC↓AAT	TGAGAAA
DGCC7710$_{\Phi2972}^{+S7}$	S7 (30 nt)	Phage 2972 (10299)	-	AAGCAAGTTGATATATTTCTCTTTCTT↓TAT	TAAGAAA
DGCC7710$_{\Phi2972}^{+S41}$	S41 (30 nt)	Phage 2972 (31518)	-	TTCCCTTCGATAATGGCAAGACCGAAA↓CGT	TCAGAAA
DGCC7710$_{\Phi2972}^{+S42}$	S42 (30 nt)	Phage 2972 (31084)	-	ATATTCATATTCCCTGCTCATGTTTGA↓TAG	CAAGAAT
DGCC7710$_{pNT1}^{+S46}$	S46 (30 nt)	Plasmid pNT1 (528)	-	TTTCCCAATCTTCTGGAATTGAATCGG↓GAT	AGAG**T**A**G**

[a] The mismatches with the consensus PAM (NNAGAAW) are underlined.

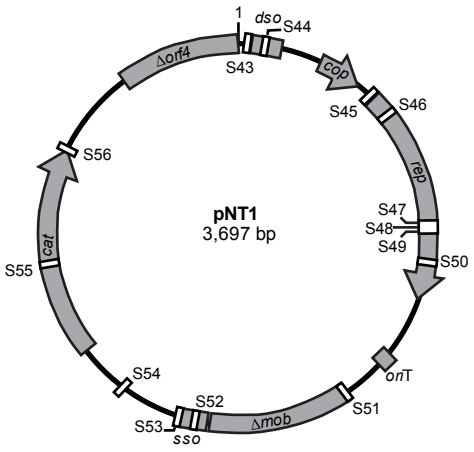
FIGURES LEGENDS

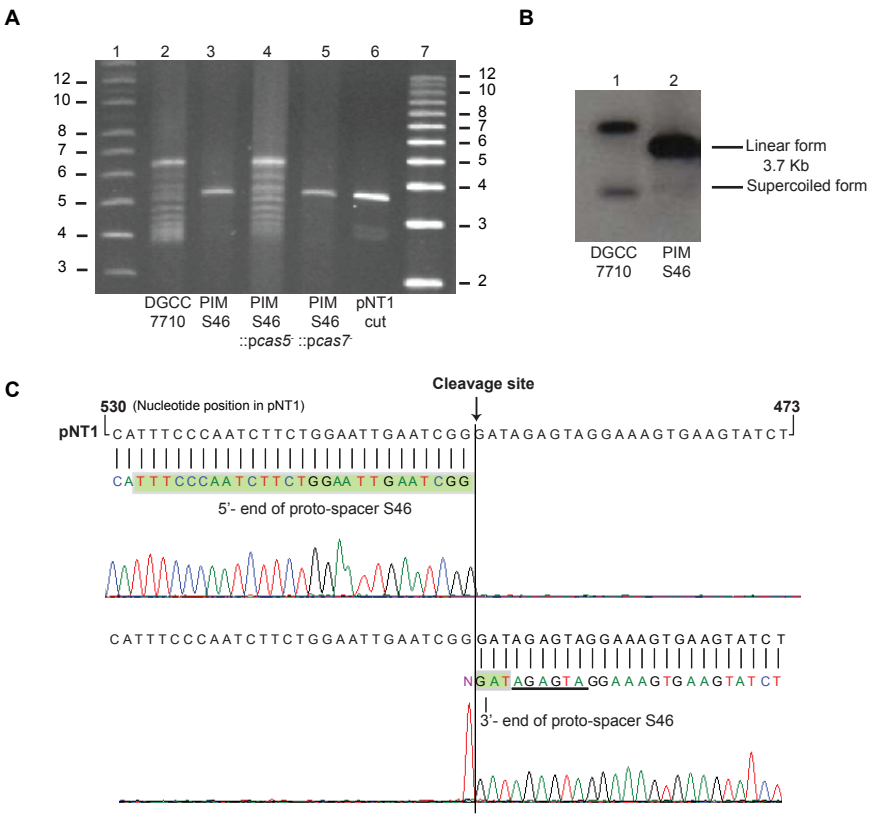
Figure 1. pNT1 proto-spacers. The numbers outside and inside the map correspond to proto-spacers from the positive and negative strand, respectively.

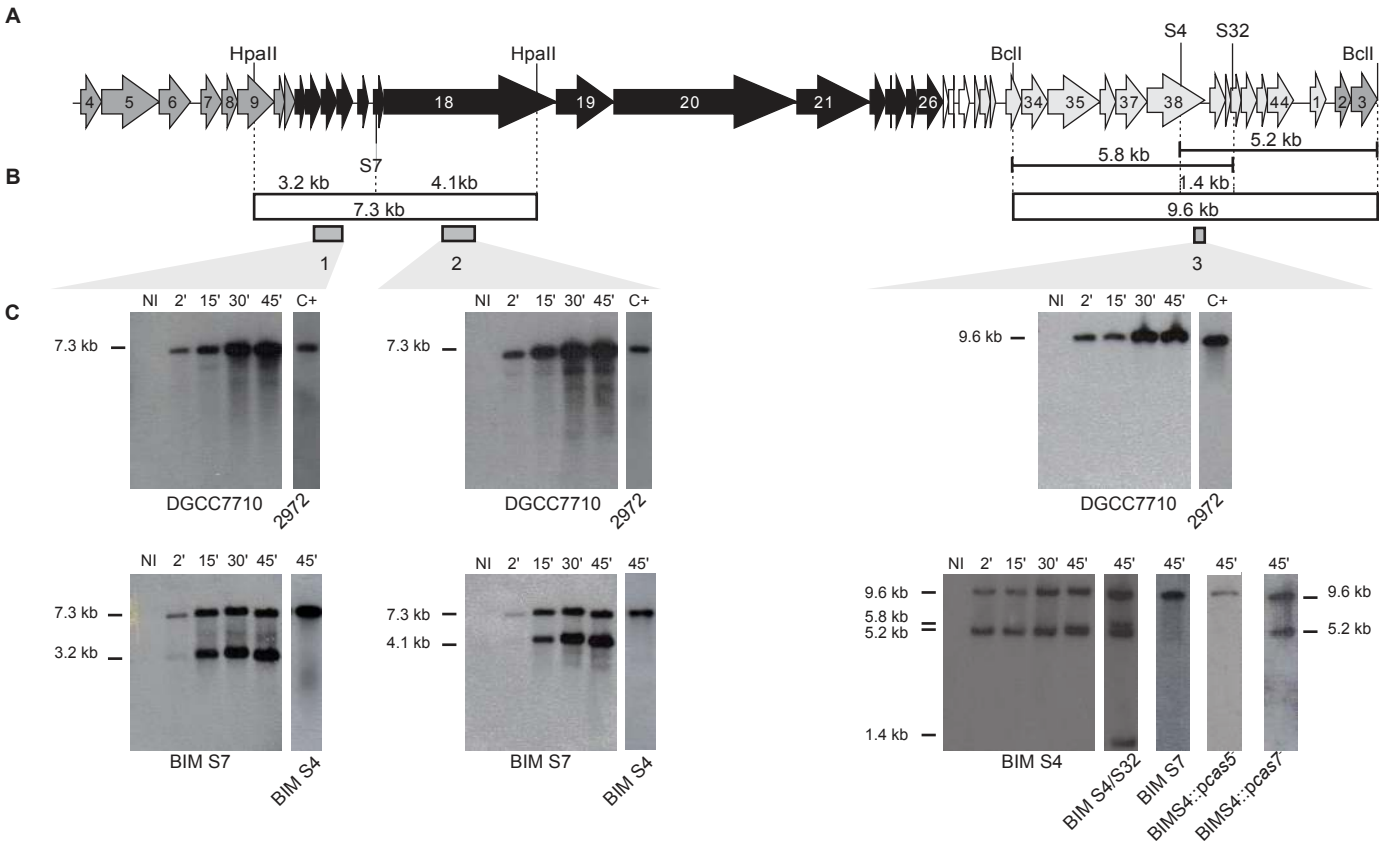
Figure 2. The CRISPR1/Cas system in *S. thermophilus* targets incoming plasmid DNA. Panel A. Plasmid pNT1 in *S. thermophilus* strains. Twenty ng of plasmid DNA were loaded per lane. pNT1 plasmid was extracted from wild-type strain DGCC7710 (lane 2), from plasmid-interfering mutant (PIM) strain S46 (lane 3), PIM S46::p*cas5*$^-$ (lane 4), and PIMS46::p*cas7*$^-$ (lane 5). pNT1 from lane 2 was linearized with EcoRV (lane 6). Lanes 1 and 7: supercoiled and 1 kb DNA ladders (Invitrogen), respectively. Panel B. Southern hybridization of plasmid pNT1 in *S. thermophilus* strains. Lane 1. One ng of native pNT1 plasmid extracted from wild-type strain DGCC7710. Lane 2: 10 ng of pNT1 extracted from PIM S46. The DIG-labeled probe 6 (Sup. Table 2) was used. Panel C. Direct sequencing electropherograms from primers NT17215_3 (upper part) and NT17225_2 (lower part) (see full methods information). The non-templated addition of adenine (T in the reverse complement sequence shown here) at the extremity of the primer NT17225_2 sequence is a sequencing artifact due to the polymerase35.

Figure 3. The CRISPR1/Cas system targets phage DNA. Panel A. Phage 2972 genome and position of proto-spacers. Light grey, dark grey, and black arrows indicate early, middle, and late transcription module, respectively[28]. Proto-spacers above and below the genome indicate positive and negative strand, respectively. Panel B. Restriction fragments detected by hybridization

and position of probes (1, 2, and 3). Panel C. Southern blots of phage-infected strains DGCC7710, BIM S4, BIM S7, BIM S4/S32, BIM S4::p*cas5*$^{-}$, and BIM S4::p*cas7*$^{-}$ over 45 minutes. NI: Non-infected strain. C+, positive control, 10 ng of digested DNA from phage 2972. One µg of total DNA from DGCC7710 and 10 µg of BIM per lane.


1
dso
S44
S43
Δorf4
copo
S45
S46
rep
S47
S48
S49
S50
pNT1
3,697 bp
S56
cat
S55
S54
S52
S53
sso
Δmob
S51
oriT


A
1 2 3 4 5 6 7
12
10
8
7
6
5
4
3
DGCC 7710
PIM S46
PIM S46 ::pcas5-
PIM S46 ::pcas7-
pNT1 cut
B
1 2
Linear form
3.7 Kb
Supercoiled form
DGCC 7710
PIM S46
C
Cleavage site
530 (Nucleotide position in pNT1)
473
pNT1
CATTTCCCAATCTTCTGGAATTGAATCGGGATAGAGTAGGAAAGTGAAGTATCT
5'- end of proto-spacer S46
3'- end of proto-spacer S46


A
HpaII
HpaII
BclI
S4
S32
BclI
4
5
6
7
8
9
18
19
20
21
26
34
35
37
38
44
1
2
3
S7
5.8 kb
5.2 kb
B
3.2 kb
4.1kb
7.3 kb
1.4 kb
9.6 kb
1
2
3
C
NI 2' 15' 30' 45' C+
7.3 kb
DGCC7710
2972
NI 2' 15' 30' 45' C+
7.3 kb
DGCC7710
2972
NI 2' 15' 30' 45' C+
9.6 kb
DGCC7710
2972
NI 2' 15' 30' 45' 45'
7.3 kb
3.2 kb
BIM S7
BIM S4
NI 2' 15' 30' 45' 45'
7.3 kb
4.1 kb
BIM S7
BIM S4
NI 2' 15' 30' 45' 45' 45' 45' 45'
9.6 kb
5.8 kb
5.2 kb
1.4 kb
9.6 kb
5.2 kb
BIM S4
BIM S4/S32
BIM S7
BIMS4::pcas5
BIMS4::pcas7