6206–6217 *Nucleic Acids Research, 2018, Vol. 46, No. 12* *Published online 28 May 2018*
doi: 10.1093/nar/gky460

Activation of the mismatch-specific endonuclease EndoMS/NucS by the replication clamp is required for high fidelity DNA replication

Sonoko Ishino[1,*], Stéphane Skouloubris[2,3], Hanae Kudo[1], Caroline l'Hermitte-Stead[3], Asmae Es-Sadik[3], Jean-Christophe Lambry[3], Yoshizumi Ishino[1] and Hannu Myllykallio[3,*]

[1]Department of Bioscience and Biotechnology, Graduate School of Bioresource and Bioenvironmental Sciences, Kyushu University, Fukuoka 8128581, Japan, [2]Department of Biology, Univ. Paris-Sud, Univ. Paris-Saclay, Orsay F-91405, France and [3]Laboratory of Optics and Biosciences, CNRS-INSERM-Ecole Polytechnique, 91128 Palaiseau France

Received March 23, 2018; Revised May 09, 2018; Editorial Decision May 13, 2018; Accepted May 14, 2018

ABSTRACT

The mismatch repair (MMR) system, exemplified by the MutS/MutL proteins, is widespread in Bacteria and Eukarya. However, molecular mechanisms how numerous archaea and bacteria lacking the *mutS*/*mutL* genes maintain high replication fidelity and genome stability have remained elusive. EndoMS is a recently discovered hyperthermophilic mismatch-specific endonuclease encoded by *nucS* in *Thermococcales*. We deleted the *nucS* from the actinobacterium *Corynebacterium glutamicum* and demonstrated a drastic increase of spontaneous transition mutations in the *nucS* deletion strain. The observed spectra of these mutations were consistent with the enzymatic properties of EndoMS *in vitro*. The robust mismatch-specific endonuclease activity was detected with the purified *C. glutamicum* EndoMS protein but only in the presence of the β-clamp (DnaN). Our biochemical and genetic data suggest that the frequently occurring G/T mismatch is efficiently repaired by the bacterial EndoMS-β−clamp complex formed via a carboxy-terminal sequence motif of EndoMS proteins. Our study thus has great implications for understanding how the activity of the novel MMR system is coordinated with the replisome and provides new mechanistic insight into genetic diversity and mutational patterns in industrially and clinically (e.g. *Mycobacteria*) important archaeal and bacterial phyla previously thought to be devoid of the MMR system.

INTRODUCTION

In all three domains of life, the fidelity of DNA replication is crucial for faithful transfer of genetic information between generations, as uncorrected errors may lead to mutations potentially causing cell death, cancer and neurogenerative diseases (1). The fidelity of DNA replication is ensured by several distinct mechanisms including the capacity of DNA polymerases to select the correct nucleotide to be incorporated by the replisome (2). During the replication process, errors occasionally occur on the nascent (daughter) strand, and are detected and corrected by either the proofreading activity of DNA polymerases or the post-replicative mismatch repair (MMR) system (3). The functional role of the core MMR proteins MutL and MutS in the maintenance of the replication fidelity is exemplified by the fact that the genetic inactivation of the MMR system typically lowers the replication fidelity by two or three orders of magnitude and results in the accumulation of transition mutations. MMR also acts on mismatches and strand misalignments during recombination and controls recombination between similar, but non-identical sequences (4). Consequently, MMR inhibits interspecies recombination and gene exchange, and thus has wide implications for understanding the evolution of species (5).

In *Escherichia coli* and some other γ-proteobacteria, the ATP-dependent assembly of the MutS and MutL heteroduplex at the sites of mismatched bases leads to the activation of the MutH endonuclease activity, which selectively cleaves transiently non-methylated *dam* methylation sites (GATC), thus facilitating the selective unwinding and degradation of the mismatch by different exonucleases (6). The resulting daughter strand 'nick' created by MutH thus directs MMR to the neosynthesized strand (methyl-directed MMR). In *E. coli*, MutS and MutL strongly interact with the replication clamp (β-subunit) and the catalytic α-subunit of the

*To whom correspondence should be addressed. Tel: +33169335010; Fax: +33169335084; Email: hannu.myllykallio@polytechnique.edu
Correspondence may also be addressed to Sonoko Ishino. Tel: +81926424218; Fax: +81926423085; Email: sonoko@agr.kyushu-u.ac.jp

DNA polymerase III (7), indicating that MMR and DNA replication are intimately coupled. In most other species, including humans, for which the mismatch repair systems have been characterized, hemimethylation does not play a role in strand discrimination (8). In these cases, the strand specificity signal have been proposed to correspond to DNA strand discontinuities created by Okazaki fragment processing and/or processing of frequently misincorporated ribonucleotides (8). These daughter strand 'nicks' may facilitate the asymmetric loading of the MutSα-MutLα complex mediated by the replication clamp PCNA (9), resulting in strand-specific MMR due to the selective incision made by MutL in the vicinity of the mismatch (nick-directed MMR). Therefore, at least two independent solutions coordinated by replication clamps have emerged during evolution to differentiate the newly synthetized DNA strand with the replication errors (mismatches) from the template strand with the correct genetic information.

Previous bioinformatics studies failed to identify the genes encoding the canonical MMR proteins in the majority of archaeal species and many actinobacteria (10), thus raising the question of how high fidelity DNA replication in these species is achieved. Recent biochemical (11) and structural (12) studies revealed a mismatch-specific endonuclease activity for the archaeal EndoMS/NucS family proteins, which use a dual base flipping mechanism for mismatch recognition (13). These proteins were first identified as PCNA-interacting endonuclease acting on branched DNA substrates (14–16). Quite recently, a bacterial homolog of these archaeal proteins from *Mycobacterium smegmatis* has been implicated in mutation avoidance through genetic studies, but in contrast to the archaeal proteins, the purified and correctly folded *M. smegmatis* EndoMS/NucS lacked nuclease activity with substrates bearing a single mismatch within a duplex DNA (17).

Here we report that the purified *Corynebacterium glutamicum* EndoMS/NucS (*Cgl*EndoMS) homodimers showed a mismatch-specific endonuclease activity that was markedly stimulated by the β-clamp encoded by *dnaN in vitro*. The C-terminal five residues of the *Cgl*EndoMS were found to be essential for the β-clamp interaction and were required for mutation avoidance activity. The replication clamp may thus direct the EndoMS to mismatches that are left unrepaired by the replisome. Our study provides new mechanistic insight into coordination of the novel mismatch repair activity and causally links the biochemical and cellular DNA repair activities of EndoMS proteins. These data provide the biochemical basis for understanding the rate and the spectrum of spontaneous mutations in archaea, actinobacteria and other bacterial phyla, which were previously thought to lack the MMR system.

MATERIALS AND METHODS

Bacterial strains and growth conditions

The bacterial strains and plasmids used in this study are listed in Table 1. *E. coli* strains XL1-Blue and XL10-Gold were used as hosts for cloning experiments. *E. coli* strains JM109 and BL21 CodonPlus (DE3)-RIL were used for the overproduction of His-tagged EndoMS/NucS and β-clamp proteins, respectively. *E. coli* strains were grown at 37°C on solid or liquid Luria-Bertani media. *C. glutamicum* strains were grown at 30°C on solid or liquid brain heart infusion (BHI) media (Oxoïd). Antibiotics for the selection of recombinant *E. coli* strains were used at the following final concentrations: ampicillin (Amp), 100 μg/ml; kanamycin (Kan), 20 μg/ml; chloramphenicol (Cam), 30 μg/ml and tetracycline (Tet), 10 μg/ml. For the selection of *C. glutamicum* transformants, kanamycin and chloramphenicol were added at concentrations of 20 and 10 μg/ml, respectively.

Construction of plasmids and *C. glutamicum* mutants

The *nucS* gene of the *C. glutamicum* ATCC 13032 strain was deleted by allelic exchange using the non-polar kanamycin resistance cassette *aphA-3* (867 pb), subcloned from pUC18K (Table 1). First, this cassette was inserted between the EcoRI and BamHI sites of pMCS5, which is unable to replicate in *C. glutamicum*. Then, the 500 bp fragment located upstream from *nucS* and the 510 bp fragment located downstream from *nucS* were PCR-amplified with the primer pairs SK94/SK95 and SK96/SK97 (Supplementary Table S1), respectively. The two PCR products were subsequently digested by NdeI/EcoRI and BamHI/PstI and inserted into the flanking regions of the *aphA-3* cassette previously cloned into pMCS5, yielding pHM455 (Table 1) (Figure 1A). This plasmid was used as a suicide vector and introduced into *C. glutamicum* ATCC 13032 competent cells by electroporation as described previously (18). Transformants were selected on BHI plates containing kanamycin after an incubation at 30°C for 3 days. The allelic exchange was then checked by PCR using the following primer pairs: SK83/SK84, SK153/H50 (absence of *nucS*) and H17/SK154 (integration of *aphA-3* at the correct locus) (see Supplementary Table S1 and Supplementary Figure S1). A deletion of 15 nucleotides from the 3′ terminus of *nucS* was performed by PCR-amplification using the primer pair SK177/SK169 (Supplementary Table S1). The PCR product was digested by NdeI and EcoRI and subsequently used to replace the upstream region of *nucS* previously cloned into pHM455, yielding pHM476. This plasmid was then introduced into ATCC 13032 competent cells by electroporation and transformants were selected as previously described. The allelic exchange was monitored by PCR-amplification and sequencing using the following primer pairs: SK177/SK84, SK153/H50 and H17/SK154 (see Supplementary Table S1 and Supplementary Figure S1).

Complementation of the *C. glutamicum* Δ*nucS* mutant

To produce the *trans* complementation plasmid pHM473, the *nucS* gene was PCR-amplified with the primer pair SK163/SK164 and inserted between the HindIII/EcoRI sites of the non-integrative *E. coli* - *C. glutamicum* shuttle vector pXMJ19 carrying the *cat* gene (Table 1). The pHM473 plasmid was then transferred into the *C. glutamicum* Δ*nucS::aphA-3* (Kan^R) strain by electroporation. Transformants were selected on BHI plates containing chloramphenicol (10 μg/ml) and kanamycin (20 μg/ml) after an incubation at 30°C for 2 days.

Table 1. Bacterial strains and plasmids used in this study

Strains or plasmids	Genotype/characteristics	Reference or origin
***E. coli* strains**		
XL1-Blue	*recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac* [F′ *proAB lacI*^{q}*ZΔM15* Tn10 (Tet^r)]	Agilent
XL10-Gold	Tet^r *Δ(mcrA)183 Δ(mcrCB-hsdSMR-mrr)173 endA1 supE44 thi-1 recA1 gyrA96 relA1 lac* Hte [F′ *proAB lacIqZΔM15* Tn10 (Tet^r) Amy Cam^r]	Agilent
JM109	*endA1, recA1, gyrA96, thi, hsdR17 (rk$^-$, mk$^+$), relA1, supE44, Δ(lac-proAB)*, [F′ *traD36, proAB, laqI*^{q}*ZΔM15*].	Takara Bio
BL21 CodonPlus (DE3)-RIL	*E. coli B F$^-$ ompT hsdS(rBmB) dcm+ Tetr gal λ(DE3) endA* Hte *[argU ileY leuW* Cam^r *]*	Agilent
***C. glutamicum* strains**		
Cg ATCC 13032	Wild-type parental strain	(20)
Cg (Δ*nucS*)	*Cg* 13032 (Δ*nucS*::*aphA-3*) *(replacement of nucS by a non-polar kanamycin resistance cassette (aphA-3))*	This work
Cg (*nucS*ΔCter)	*Cg* 13032 (*nucS*ΔCter::*aphA-3*) *(replacement of nucS by a copy of nucS with its 15 last nucleotides deleted, followed by a non-polar kanamycin resistance cassette)*	This work
Cg (Δ*nucS*) + pXMJ19	*Cg* 13032 (Δ*nucS*::*aphA-3*) transformed with the empty pXMJ19 vector	This work
Cg (Δ*nucS*) + pHM473	*Cg* 13032 (Δ*nucS*::*aphA-3*) transformed with the pHM473 plasmid carrying a wild-type copy of the *nucS* gene	This work
Plasmids		
pMCS5	*ori pMB1 (colE1), plac, lacZα*, Amp^r, MCS of pUC18 / Unable to replicate in *C. glutamicum* (suicide vector)	MoBiTec
pXMJ19	*ori pMB1 (colE1), ptac, lacI*^{q}, Cam^r, MCS of pUC18 / *E. coli-C. glutamicum* (shuttle vector)	(37)
pQE80L	*ori colE1, pT5, lacI*^{q}, Amp^r, 6xHis tag (N-ter)	Qiagen
pG-KJE8	*ori p15A, paraB, dnaK, dnaJ, grpE,* Cam^r	Takara Bio
pUC18K	*ori pMB1 (colE1), plac, lacZα*, Amp^r, Kan^r / Promotor less non-polar kanamycin cassette (*aphA-3*) cloned into pUC18	(38)
pHM449	*nucS* of *Cg* ATCC 13032 cloned into pQE80L (BamHI / PstI)	This work
pHM455	*aphA-3* cassette flanked with the upstream and downstream regions of *nucS* and cloned into pMCS5 (Amp^r,Kan^r)	This work
pHM473	*nucS* of *Cg* ATCC 13032 cloned into pXMJ19 (HindIII / EcoRI)	This work
pHM476	*nucS* with its 15 last nucleotides deleted followed by the *aphA-3* cassette and cloned into pMCS5 (Amp^r,Kan^r)	This work
pHM477	A stop codon was introduced into *nucS* in pHM449 to delete five amino acids from the C-terminus of *Cgl*EndoMS	This work
pET-*Cgl*β-clamp	*dnaN* of *Cg* ATCC 13032 cloned into a modified version of pET28a (NdeI / NotI), *ori pMB1 (colE1),pT7,* Amp^r, 6xHis tag (N-ter)	This work

Estimation of mutation frequencies

C. glutamicum and its derivatives were grown overnight using 2 ml 96-deep-well plates at 30°C in BHI liquid medium, including antibiotics when required. Appropriate dilutions were plated on BHI medium with or without rifampicin (Rif) at 100 μg/ml. Rifampicin resistant (Rif^R) colonies originating from the independent cultures obtained were isolated and sequenced for the wild-type strain 13032 and the Δ*nucS* mutant, respectively. Spontaneous mutations conferring rifampicin resistance were then identified from these mutants, by PCR-amplification and sequencing a well characterized Rif^R-determining region in the *rpoB* gene (NP_599733; NCgl0471), using the primer pair SK157/SK158 [Supplementary Table S1 (19)]. Where indicated, streptomycin was used in analogous experiments at the final concentration of 100 μg/ml.

Cloning of the *C. glutamicum* genes encoding EndoMS/NucS and β-clamp (DnaN) for protein purification

The *nucS* (CAF19919; NCgl1168) and *dnaN* (NP_599254; NCgl0002) genes, encoding the *Cgl*EndoMS/NucS and the DNA polymerase III β-subunit (β-clamp) respectively, were amplified by PCR from *C. glutamicum* ATCC 13032 genomic DNA, using *Pfu* DNA polymerase (Agilent) and the following primer pairs: SK83/SK84 for *nucS* and *Cgl*β-clamp-F/*Cgl*β-clamp-R for *dnaN* [Supplementary Table S1 (20)]. The amplified *nucS* gene was inserted between the BamHI and PstI sites of pQE80L (Qiagen), to generate pHM449 (Table 1). To delete the last five amino acids of the EndoMS/NucS protein, a stop codon was introduced into the *nucS* gene on the pHM449 plasmid by PCR-mediated mutagenesis (QuikChange site-directed mutagenesis kit; Agilent), using the primer pair ΔCter-F/ΔCter-R (Supplementary Table S1). The DNA product corresponding to *dnaN* was digested by NdeI and NotI (New England Biolabs), and inserted into the respective sites of a modified pET28a vector (Novagen). In this modified vector, the recognition sequence for thrombin was changed to that for Tobacco Etch Virus (TEV) protease, and the Kan-resistance gene was changed to an Amp-resistance gene. The resulting plasmid was designated as pET-*Cgl*β-clamp (Table 1). Both inserted sequences were confirmed by sequencing.

Protein expression and purification

His-tagged *Cgl*EndoMS and *Cgl*EndoMS-ΔCter were prepared from *E. coli* JM109 cells transformed with pHM449 and pHM477, respectively. The chaperone plasmid pG-KJE8 was co-transformed with both plasmids. *E. coli* cells

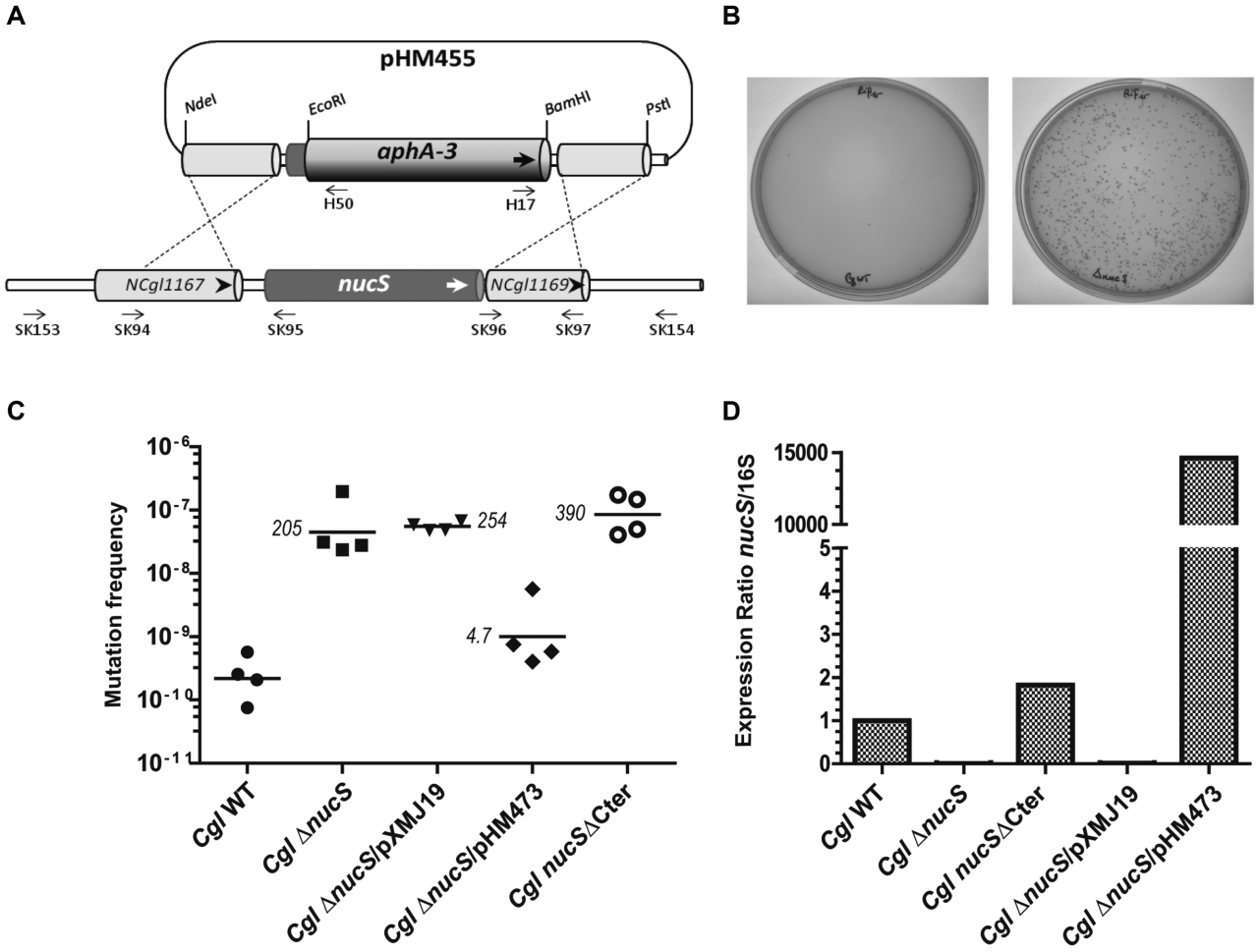

Figure 1. The functional role of *Cgl*EndoMS/NucS in mutation avoidance. (**A**) The physical and genetic organization of the *C. glutamicum nucS* locus are shown. Approximate positions of the oligonucleotides used for construction and confirmation of the Δ*nucS::aphA-3* (Kan^R) are also indicated. (**B**) The plates showing the formation of spontaneous Rif^R colonies for wildtype (left panel) and Δ*nucS::aphA-3* (right panel) strains. The plates demonstrate the significant increase in the spontaneous mutation rate for the deletion strain. (**C**) Mutation frequencies observed for the *C. glutamicum* wild type, Δ*nucS*, Δ*nucS*/pXMJ19 (empty plasmid control), Δ*nucS*/pHM473 (carrying *nucS in trans*) and *nucS*ΔCter strains. The last strain lacks the five last amino acid residues of *Cgl*EndoMS/NucS. The line indicates the geometric mean of the four different biological replicates. Numbers refer to relative fold differences compared to *Cgl*EndoMS-WT. (**D**) Relative *nucS* expression levels for the different strains are shown. The values shown are averages of two biological replicates determined using RT-PCR as described in the materials and methods section.

were cultured in LB medium, containing 50 μg/ml ampicillin, 34 μg/ml chloramphenicol, 10 ng/ml tetracycline and 0.5 mg/ml L-arabinose, at 37°C until the culture attained an OD_{600} of 0.5. IPTG was then added to a final concentration of 0.2 mM, and the cells were further grown for 18 h at 25°C. The cells were collected, and were resuspended in buffer A (25 mM HEPES, pH 7.5, 1 M NaCl, and 5 mM imidazole), and were disrupted by sonication. The soluble cell extracts were applied to TALON metal affinity resin (Clontech), and the resin was washed with buffer A. The bound fraction was eluted with 25 mM HEPES, pH 7.5, 0.5 M NaCl and 100 mM imidazole, and the eluate was applied to Superdex 200 5/150 column (GE Healthcare), which was eluted with 25 mM HEPES, pH 7.5, 0.3 M NaCl, 0.5 mM DTT, and 0.1 mM EDTA. The fraction containing *Cgl*EndoMS was stored at –80°C. His-tagged *Cgl*β-clamp was prepared from *E. coli* BL21 CodonPlus (DE3)-RIL (Agilent) cells transformed with pET-*Cgl*β-clamp. *E. coli* cells were cultured in LB medium, containing 50 μg/ml ampicillin and 34 μg/ml chloramphenicol, at 37°C until the culture attained an OD_{600} of 0.5. IPTG was then added to a final concentration of 1 mM, and the cells were further grown for 18 h at 25°C. The cells were collected, and were resuspended in buffer B (25 mM HEPES, pH 7.5, 0.5 M NaCl and 5 mM imidazole), and were disrupted by sonication. The soluble cell extracts were applied to TALON metal affinity resin, and the resin was washed with buffer B. The bound fraction was eluted with 25 mM HEPES, pH 7.5, 0.3 M NaCl and 100 mM imidazole, and the eluate was applied to a Superdex 200 5/150 column, which was eluted with 25 mM HEPES, pH 7.5, 0.3 M NaCl, 0.5 mM DTT and 0.1 mM EDTA. The protein concentrations were calculated by measuring the absorbance at 280 nm. The theoretical molar extinction coefficients of *Cgl*EndoMS, *Cgl*EndoMS-ΔCter, and *Cgl*β-clamp are 19 940, 19 940 and 21 430 M^{-1} cm^{-1},

respectively. The purity of the proteins was confirmed by SDS-PAGE.

Gel filtration chromatography analyses

Purified recombinant *Cgl*EndoMS (1.2 nmol), *Cgl*EndoMS-ΔCter (1.2 nmol), *Cgl*β-clamp (2 nmol) were applied to a Superdex 200 3.2/30 column (GE Healthcare), and were eluted at a flow rate of 40 μl/min in buffer containing 25 mM HEPES, pH 7.5, 0.5 mM DTT, 0.1 mM EDTA and 0.15 M NaCl using SMART system (Amersham Pharmacia). The standard marker proteins, including thyroglobulin (MW, 670 000), γ-globulin (MW, 158 000), ovalbumin (MW, 44 000) and myoglobin (MW, 17 000), were subjected to gel filtration as a control.

Cleavage assay of mismatch-containing DNA

Cleavage reactions of the Cy5-labeled substrate dsDNAs, formed with the oligonucleotides listed in Supplementary Tables S2 and S3, by the proteins were generally performed at 30˚C in an optimized reaction mixture containing 25 mM Bis–Tris, pH 6.4, 2.5 mM $MnCl_2$, 5 mM DTT and 0.1% Triton X-100. The modified reaction conditions are described in each section. The reaction was quenched by the addition of 25 mM EDTA and 0.3% SDS, and the products were analyzed by either 10% native polyacrylamide gel electrophoresis (PAGE) or 8 M urea–15% PAGE in TBE buffer (89 mM Tris, 89 mM boric acid and 2.5 mM EDTA, pH 8.3). The products were visualized with a Typhoon Trio+ image analyzer (GE Healthcare).

Electrophoretic mobility shift assay

*Cgl*EndoMS (0.8 μM as a monomer) and *Cgl*β-clamp (8 μM as a monomer) were incubated with 5 nM 5′-Cy5-labeled ssDNA (45 nt) or dsDNA (45 bp) in a reaction solution (25 mM Bis–Tris, pH 6.4, 2.5 mM $MnCl_2$, 5 mM DTT and 0.1% Triton X-100) at 30˚C for 30 min. An aliquot was fractionated on an 8% PAGE in TBE buffer and was visualized and quantified with a Typhoon Trio+ image analyzer.

Homology modeling of the bacterial EndoMS/NucS

We constructed a model of the *C. glutamicum* EndoMS/NucS dimer with the modelling program Modeller (21) version 9.18 by using the EndoMS-dsDNA structure carrying a mismatch (PDB code 5GKE) as the template. Five initial models were constructed and the best model was chosen using molpdf and the DOPE score and by evaluating the stereochemical quality with PROCHECK (22). Molecular graphics and structures superposition were performed with the UCSF Chimera package (23). Chimera was developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco (supported by NIGMS P41-GM103311).

RNA extraction, reverse transcription and real-time PCR quantification

Total RNA was extracted and purified from the *C. glutamicum* wild type strain and the mutants carrying a deletion of 15 nucleotides from the 3′ terminus of *nucS* using a Qiagen miRNeasy Mini kit (Qiagen, Courtaboeuf, France) with a final elution volume of 50 μl. The concentration of the RNA fraction was quantified using a Nanodrop 2000c system (Thermo Fisher Scientific). The RT-PCR was performed with a Maxima First Strand cDNA Synthesis Kit for RT-PCR (Thermo Scientific). For each strain, 200 ng of total RNA was used for the synthesis of the first-strand complementary DNA (cDNA). We used 50 pg of cDNA for the real-time PCR assay with SensiFast™ SYBR No-ROX reagents (Bioline Reagent) and the MiniOpticon System (BioRad). The real-time PCR was performed using a concentration range of *nucS* and 16S rRNA (house-hold) genes to calculate the quantitative expression ratio of *nucS* to 16S rRNA in each sample from two separate experiments. The following primer pairs were used: qPCR-NucS-F/qPCR-NucS-R (*nucS*) and qPCR-16S-F/qPCR-16S-R (16S rRNA; NCg1r01) (Supplementary Table S1). Where indicated, the percentage of the expression ratio of *nucS* to 16S for the wild type sample was set to 100%.

RESULTS

Identification and preparation of a *C. glutamicum* EndoMS

Previous structural studies have demonstrated that the EndoMS proteins possess a two-domain structure (12,16). The amino-terminal dimerization domain has been implicated in the mismatch (MM) binding, whereas the carboxyl-terminal region bears an endonuclease active site, which is structurally related to many families of restriction endonucleases (11,12). As we previously reported, the EndoMS homologs were identified not only in archaeal, but also in some of bacterial phyla, particularly actinobacteria (11,16). EndoMS also reportedly interacts with proliferating cell nuclear antigen (PCNA) via a PCNA-interacting peptide (PIP) motif at its C-terminal region, although the PIP motif is not conserved in all homologs (11,12). The β-clamp, identified as the β subunit of DNA polymerase III, in Bacteria has the same functions as the eukaryotic and archaeal PCNAs. The sequence, [EK]-[LY]-[TR]-L-F, which is similar to the consensus sequences of the β-clamp binding motif (24) was found in the C-terminal region of the putative sequences of the actinobacterial EndoMS homologs (Supplementary Figure S1A). Among the bacterial sequences, we selected the *C. glutamicum* EndoMS (encoded by *nucS*) and the β-clamp (encoded by *dnaN*) homologs for further studies. *Cgl*EndoMS shares only 29.2%, 26.5% and 20.7% amino acid identities, with the EndoMS proteins from *Thermococcus kodakarensis*, *Pyrococcus abyssi* and *Methanopyrus kandleri,* respectively, but the critical residues for the mismatch binding and nuclease activity are highly conserved among them.

The functional role of *Cgl*EndoMS in mutation avoidance

The *nucS* gene is surrounded by the genes NCgl1167 (upstream) and NCgl1169 (downstream) genes, which have been annotated as ‘universal stress protein UspA’ and ‘nucleoside-diphosphate-sugar epimerase’ in the *C. glutamimum* genome, respectively (Figure 1A). These proteins have

predicted functions in multiple stress responses (including oxidative stress) and the metabolism of nucleotide sugars that act as activated donors in glycosylation reactions. Our strategy to inactivate the *nucS* gene, by replacing it with the *aphA-3* gene conferring kanamycin resistance (Kan^R), is also illustrated in Figure 1A. The obtained plasmid-borne deletion-insertion construct (Δ*nucS*::*aphA-3*) was inserted into the suicide plasmid to yield the plasmid pHM455 (Table 1), which was transferred to *C. glutamicum* by electroporation. We screened twenty Kan^R colonies by PCR for the absence of *nucS* resulting from an expected double recombination event. This led to the identification of one *C. glutamicum* Δ*nucS*::*aphA-3* candidate strain where *nucS* had been replaced with *aphA-3*. The correct chromosomal localization of *aphA-3* was confirmed through additional PCR analyses (see Supplementary Figure S1B and C). The isolation of this deletion strain allowed the subsequent testing of whether *nucS* is required for mutation avoidance by estimating the mutation frequencies of Rif^R in the wild type and Δ*nucS*::*aphA-3 C. glutamicum* strains (Figure 1B). Using four biological replicates and by comparing the geometric means of the measured mutation rates, we observed that the lack of *nucS* increased the spontaneous mutation rate by a factor of 205 in comparison to the wild type strain (Figure 1C). This hypermutable phenotype was suppressed by expressing *nucS in trans* from the plasmid pHM473 (Figure 1C and D), thus establishing the role of *nucS* in mutation avoidance. To support this result, we measured mutation frequencies using a different marker gene. The mutation frequencies observed with streptomycin were 1.37×10^{-8} and 1.47×10^{-6} for the wild type and Δ*nucS* strains, respectively. To investigate the possible functional importance of the putative β-clamp interaction motif in *Cgl*EndoMS (see Figure 2A), we deleted 15 nucleotides corresponding to five carboxy-terminal residues from the chromosomal *nucS* gene and measured the level of Rif^R colony production (Table 1). Our results revealed that although this construct was still transcribed in cells, the five carboxyterminal residues, ELTLF, are absolutely required for the mutation avoidance function of *Cgl*EndoMS in our assays (Figure 1C and D).

Purification of *Cgl*EndoMS and *Cgl*β-clamp

The recombinant His-tagged proteins, *Cgl*EndoMS, *Cgl*β-clamp, and *Cgl*EndoMS-ΔCter, among which the latter lacks five carboxyterminal residues of the wild type *Cgl*EndoMS, were successfully overproduced by cultivating the *E. coli* cells bearing the expression plasmids, pHM449, pET-*Cgl*β-clamp and pHM477 with IPTG induction (Table 1). As shown in Supplementary Figure S2, the three proteins were purified to near homogeneity. The migrations of the protein bands for His-tagged *Cgl*EndoMS and *Cgl*EndoMS-ΔCter were obviously slower than those estimated from the calculated molecular weights, 27 014.5 and 26 410.8, respectively, probably because of the open conformation of the EndoMS protein without a DNA substrate. The migration positions of the protein band for the His-tagged *Cgl*β-clamp protein corresponded well to the calculated molecular weight of 44 872.9. The gel filtration of *Cgl*EndoMS and *Cgl*EndoMS-ΔCter showed similar chromatograms. The estimated molecular weights of

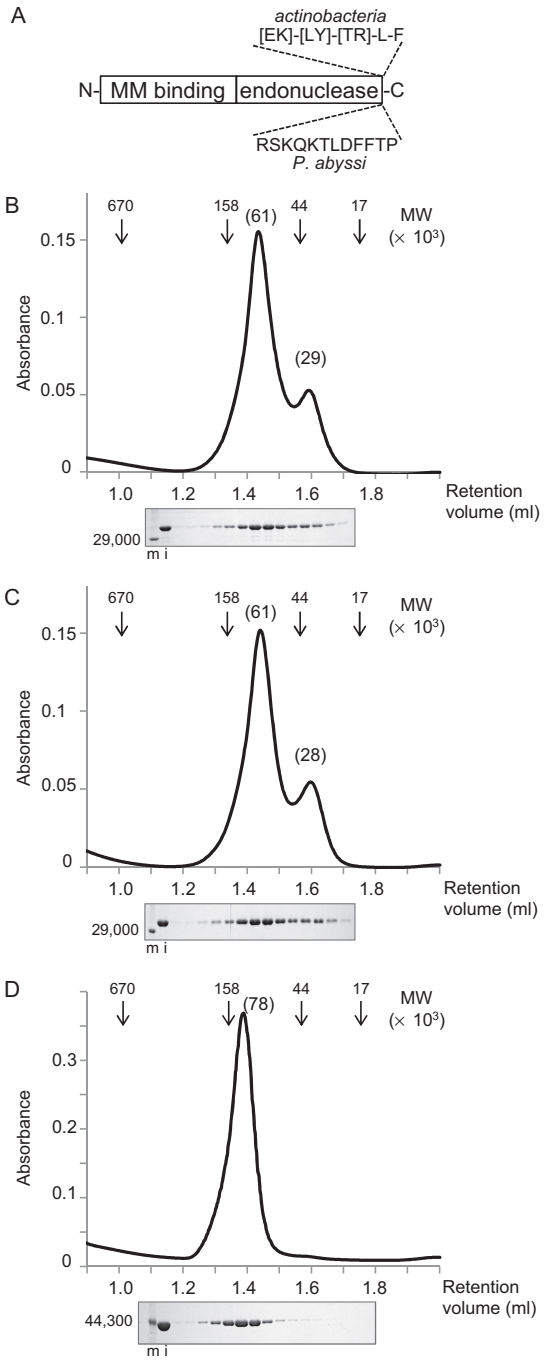

Figure 2. Properties of *Cgl*EndoMS-WT and *Cgl*β-clamp. (**A**) The two-domain structure of the EndoMS/NucS proteins. 'MM binding' and 'endonuclease' refer to mismatch binding and endonuclease active site domains, respectively. The positions of the experimentally determined and putative replication clamp binding motif for the *Thermococcales* and actinobacterial EndoMS/NucS proteins are shown. (B–D) Gel filtration chromatography analyses of *Cgl*EndoMS-WT (**B**), *Cgl*EndoMS-ΔCter (**C**), and *Cgl*β-clamp (**D**). Elution profiles, monitored by the absorbance at 280 nm, are shown in the upper part. The arrowheads indicate the elution positions of the standard marker proteins, and the numbers in parentheses on the peak tops indicate the relative molecular masses estimated by the standard curve from the marker proteins. Aliquots of each fraction were subjected to an SDS-12% PAGE followed by Coomassie Brilliant Blue staining (lower part). Protein size markers were run in lanes m, and their sizes are indicated on the left sides of the gels. The loading samples (20%) were run in lanes *i*.

the peaks were 61 000 and 29 000 or 28 000, which appeared to be a dimer and a monomer, respectively (Figure 2B and C). This result suggests that, similar to its archaeal orthologues, *Cgl*EndoMS is mainly a dimer in solution. The main peak was isolated and applied to the same column again. However, this rechromatography resulted in the same profiles with two peaks, suggesting that the dimer form of *Cgl*EndoMS is unstable in both cases of WT and ΔCter. This might also reflect the presence of the monomer-dimer equilibrium for the two *Cgl*EndoMS variants. The elution profile of the *Cgl*β-clamp showed a clear single peak, with an estimated molecular weight of ∼77 900, suggesting that it exists as a stable homodimer in solution (Figure 2D). A similar result was previously described for the *E. coli* β-clamp (25).

The mismatch-specific cleavage activity of *Cgl*EndoMS is activated by the β-clamp

To investigate whether *Cgl*EndoMS has mismatch DNA-specific endonuclease activity, the purified protein was mixed with oligodeoxynucleotides containing a mismatched base-pair, and the cleavage reactions were monitored by native PAGE. We found that in comparison to *T. kodakarensis* EndoMS (20 nM), *Cgl*EndoMS had less activity with the mismatched DNA, even when more than twice the protein concentration (50 nM) was used (Figure 3A, lanes 1 and 3). However, the cleavage activity was markedly stimulated by the addition of the *Cgl*β-clamp, and maximal stimulation was reached at equimolar amounts of the *Cgl*β-clamp and *Cgl*EndoMS (Figure 3A, lanes 3–7, and 3B). Furthermore, no stimulation was observed when the *Cgl*EndoMS-ΔCter mutant was used (lanes 8–12). SPR analysis with immobilized *Cgl*β-clamp were performed to quantify the physical interactions of *Cgl*EndoMS with *Cgl*β-clamp (Figure 3C). A K_D value of 88 nM was estimated through global fitting analysis, accounting for both association and dissociation phases using a 1:1 interaction model. The apparent binding affinity of EndoMS to β-clamp is comparable to that of *P. abyssi* EndoMS to PCNA (K_D = 15 nM) (16). On the other hand, *Cgl*EndoMS-ΔCter did not show any response with *Cgl*β-clamp even at a high concentration up to 2 μM (Figure 3C). These results indicated that the *Cgl*β-clamp has a critical function in the activation of the mismatch-specific endonuclease activity of *Cgl*EndoMS, and that the last five residues of *Cgl*EndoMS are crucial for the interaction with the *Cgl*β-clamp. To determine the exact cleavage site of *Cgl*EndoMS in the mismatch-containing DNA, synthetic DNA substrates, labeled with Cy5 at the 5′-terminus of the upper or lower strand were used as substrates. The products generated by *Cgl*EndoMS with and without the *Cgl*β-clamp were analyzed by denaturing PAGE. These experiments revealed that *Cgl*EndoMS cleaved the third phosphodiester bond on the 5′ side of the mismatched base in both strands, to produce a cohesive end with a five-nucleotide long 5′-protrusion (Figure 4). The cleavage pattern by *Cgl*EndoMS was not different with or without *Cgl*β-clamp (Figure 4B and C), and was exactly the same as that observed in the presence of *Tko*EndoMS (11). This observation is also fully compatible with our structural models of the *Cgl*EndoMS dimer using the *Tko*EndoMS-dsDNA structure (PDB code

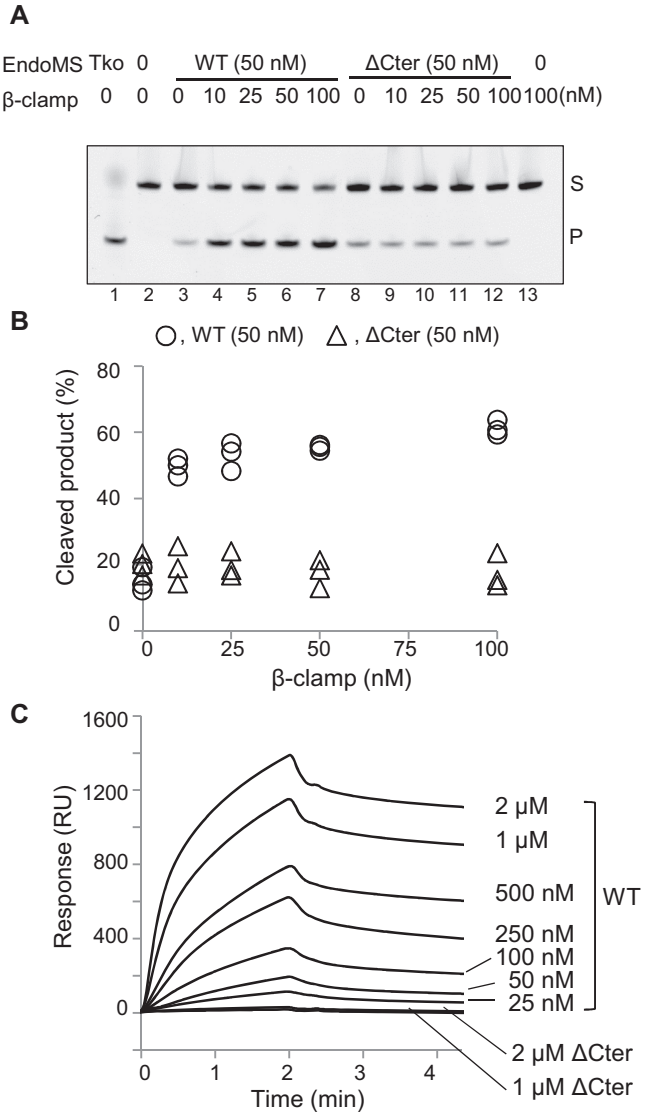

Figure 3. Cleavage of mismatch-containing DNA by *Cgl*EndoMS-WT and *Cgl*EndoMS-ΔCter with *Cgl*β-clamp. (**A**) The 5′-Cy5-labeled DNA substrates (5 nM) containing the GT base pair were incubated with the proteins. The products were analyzed by native 10% PAGE followed by laser scanning (cropped gel image). Lanes 1, positive control (20 nM TkoEndoMS, as a monomer); 2, negative control (no proteins); 3–7, 50 nM *Cgl*EndoMS-WT (as a monomer); 8–12, 50 nM *Cgl*EndoMS-ΔCter (as a monomer); 13, no EndoMS. The indicated concentrations (as a monomer) of *Cgl*β-clamp were added to the reactions. Representative results are shown. The band assignments are indicated on the side of the panels: s, substrates: p, cleaved products. (**B**) Quantification of the cleaved products. Independent data points from three measurements are plotted. (**C**) The physical interactions of *Cgl*EndoMS with *Cgl*β-clamp were characterized by SPR using a BIACORE J system (Biacore Inc.). Purified *Cgl*β-Clamp was immobilized on the sensor Chip CM5, and various concentrations (indicated on the right sides of the sensorgrams) of *Cgl*EndoMS-WT and *Cgl*EndoMS-ΔCter were loaded onto the chip for 2 min at a flow rate of 30 μl/min in running buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 3 mM EDTA, and 0.05% surfactant P20). Regeneration was achieved using a 1-min injection of 0.025% SDS. Each background response was subtracted. Data were analyzed using BIAevaluation v.3 software (Biacore Inc.)

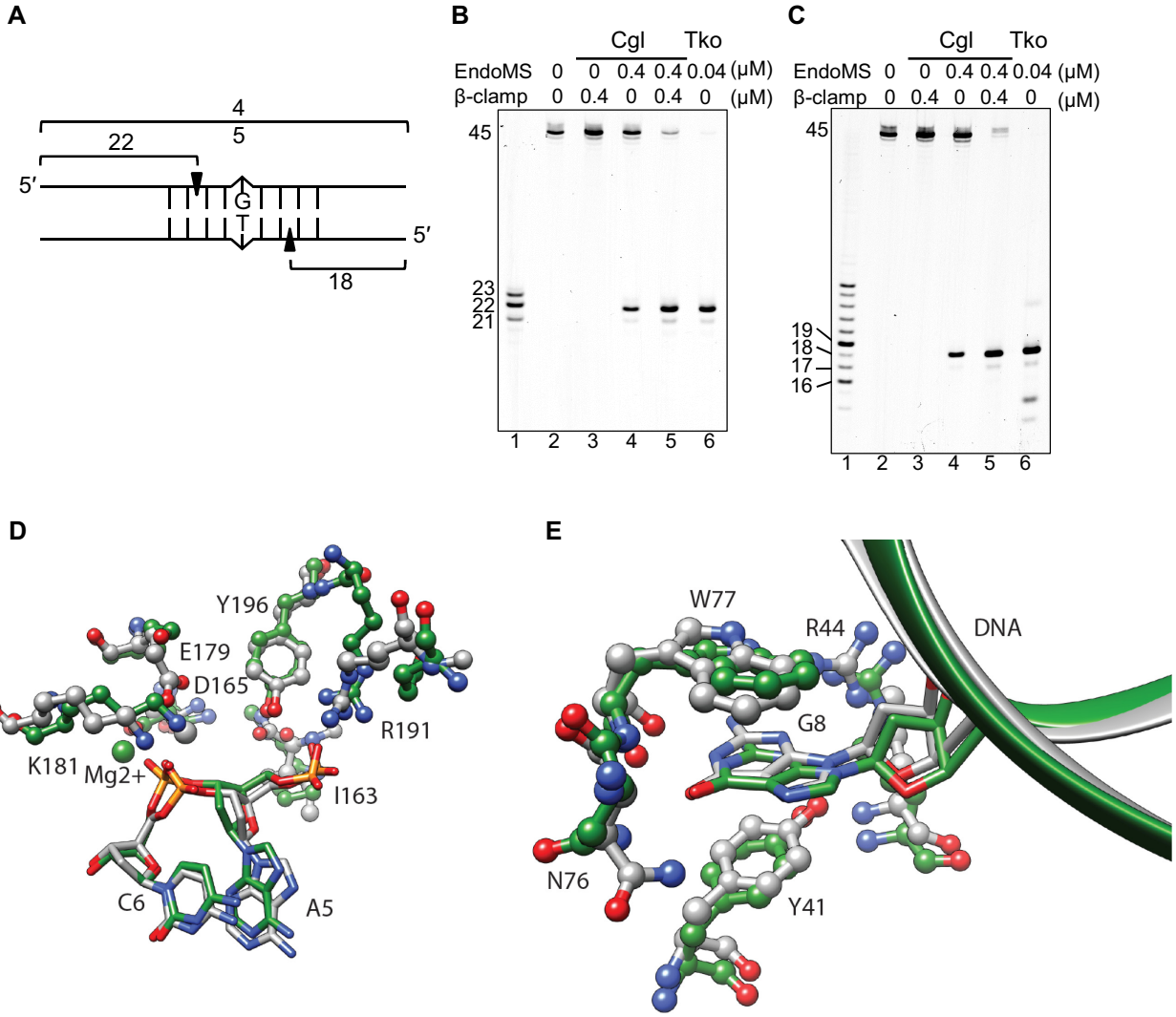

Figure 4. Cleavage pattern of mismatch-containing DNA by *Cgl*EndoMS. (**A**) The dsDNA substrate (45 bp) containing the mismatched base pair and the cleaved products are shown schematically. The 5′-Cy5-labeled upper strand (B) and the 5′-Cy5-labeled lower strand (C) were used to make the dsDNA. The numbers indicate the length of the each strand cleaved by *Cgl*EndoMS. (**B** and **C**) The substrates (5 nM) were incubated with various proteins. The products were separated by 8 M urea–12% PAGE in TBE buffer followed by laser scanning. The size markers were loaded in lanes 1 in panels b and c, and the sizes are indicated on the left side of each band. Lanes 2, no proteins; 3, 0.4 μM *Cgl*β-clamp (as monomer); 4, 0.4 μM *Cgl*EndoMS (as monomer); 5, 0.4 μM *Cgl*EndoMS (as monomer) and 0.4 μM *Cgl*β-clamp (as monomer); 6, 0.04 μM *Tko*EndoMS (as monomer). (**D**) Catalytic site structure superposition of EndoMS/NucS model (green) and template 5GKE (gray). DNA and protein are depicted in stick and ball-and-stick representations respectively. Residue numbering is relative to the template sequence. R191 is not a conserved amino acid, but in the case of *Cgl*EndoMS/NucS R173 is predicted to present a similar interaction with DNA phosphate. (**E**) Mismatch base recognition site structure superposition of EndoMS/NucS model (green) and template 5GKE (gray). The DNA and protein are depicted in stick-and-ribbon and ball-and-stick representations, respectively. Residue numbering is relative to the template sequence.

5GKE), indicating the excellent superimposition of the catalytic and substrate binding sites (Figure 4D and E).

DNA binding ability of *Cgl*EndoMS

A recent report revealed that the EndoMS/NucS protein from *M. smegmatis* has some ssDNA binding ability, although no cleavage activity for mismatched DNA was shown *in vitro* (17). For comparison with the *M. smegmatis* EndoMS/NucS, the DNA binding ability of *Cgl*EndoMS was examined by an electrophoresis mobility shift assay (EMSA). The EMSA revealed that *Cgl*EndoMS cleaved the G/T-containing DNA and no shifted band was observed. In the case of normal dsDNA and ssDNA, some shifted bands were observed as reported for *M. smegmatis* EndoMS/NucS (Supplementary Figure S3). These shifts may reflect nonspecific binding.

Optimized reaction conditions for actinobacterial EndoMS with mismatch-containing DNA

To further characterize *Cgl*EndoMS and *Cgl*β-clamp, the cleavage reactions were performed under various conditions. The effects of divalent cations, temperature, and pH on the activity are shown in Supplementary Figure S4. Whereas the mismatch endonuclease activity of

*Tko*EndoMS can use both Mg^{2+} and Mn^{2+} for catalysis and functions in a broad pH range ranging from 6.0 to 11.0, the activity of the bacterial protein is strictly Mn^{2+}-dependent and has a very narrow pH optimum around 6.4. In agreement with the mesophilic lifestyle of *C. glutamicum*, the activity of *Cgl*EndoMS rapidly diminishes above 30°C. In the future, comparisons of bacterial and archaeal EndoMS orthologues will be interesting to further investigate the structural basis of their peculiar catalytic differences, although they do not seem to play major roles in the determination of the substrate specificity as described below.

In vitro mismatch specificity of the *Cgl*EndoMS protein

To investigate the mismatch specificity of the *Cgl*EndoMS nuclease, we used substrates containing all possible combinations of mismatched bases in cleavage reactions in the presence of the *Cgl*β-clamp (Figure 5). We observed very clear endonuclease activities not only with the G/T mismatch-containing dsDNA substrate used in the above experiments and preferred by the *Tko*EndoMS, but also with DNA duplexes containing T/T and, to a lesser extent, G/G mismatches. However, T/C, A/G, C/C and A/C mismatches were either cut very poorly or not all. We conclude that the preferred *in vitro* substrates of *Cgl*EndoMS are G/T (mismatch resulting in 'transition' mutations if left unrepaired) and T/T ('transversion' mismatch), which were cleaved with high efficiencies under these experimental conditions. Notably, the other possible 'transition mismatch', A/C, was not cleaved at all under these optimized conditions. We also stress that a previous study indicated that the EndoMS proteins lack 'nicking' activity that would cleave only one of the two strands of the DNAs with the mismatch-containing DNA substrates (11). Nevertheless, there is a possibility that *Cgl*EndoMS, in particular complexed with the DNA-bound *Cgl*β-clamp, might introduce a nick into a dsDNA. To address this point, a normal DNA or DNAs containing C/C or A/C mismatches were reacted with *Cgl*EndoMS with or without β-clamp. The reaction mixtures were fractionated by electrophoresis on a denaturing gel in parallel with a native gel. A cleaved band was detected from the G/T mismatch, but not from either the C/C or A/C mismatch, indicating that *Cgl*EndoMS did not clearly cleave even one of the two strands of the DNAs in the vicinity of these mismatches (Supplementary Figure S5).

Accumulation of transition mutations in the absence of the *Cgl*EndoMS

To investigate the mutational patterns of the *C. glutamicum* wild type and Δ*nucS* strains, we isolated and sequenced independent Rif^R mutants for the two strains (Figure 6). Many different types of mutations within the *rpoB* gene can cause resistance to rifampicin (19,26). We found that up to 98% of base substitutions in the isolated Rif^R mutants were transition mutations (Supplementary Table S4, Figure 6A). The fact that transitions are strongly favored over transversions among the sequenced base substitutions is in agreement with earlier studies on spectra of spontaneous mutations in Bacteria [see for instance (27)]. In particular, our results suggested that the relative level of the appearance of A:T to G:C transitions were favored approximately three to four-fold in the absence of EndoMS (Figure 6A). *A priori*, these transitions could result from either non-repaired C:A or G:T mismatches, as indicated schematically in Figure 6B.

DISCUSSION

Our previous study, searching for proteins involved in mismatch repair (MMR) in archaea that lack the homologs for MutS/MutL proteins, revealed an endonuclease family dubbed EndoMS that is encoded by *nucS*. This protein family cleaves both strands of double-stranded DNA carrying mismatches *in vitro*, thus generating double-strand breaks (11). In addition, a structural analysis of this archaeal homodimeric EndoMS, complexed with a double-stranded DNA carrying G/T mismatches, revealed that these mismatched bases were flipped out into binding sites located in the N-terminal domain of EndoMS. The active site for cleavage, which resembles that of the restriction endonuclease, is located in the C-terminal domain (12), and the archaeal EndoMS structures in the absence of DNA, together with earlier SAXS analyses (14), indicated the movement of the C-terminal catalytic domain brought about by the DNA binding that is required for activity. The *nucS* genes are widely distributed in the domain Archaea, but not in most Bacteria and not at all in Eukarya. It is remarkable that a *nucS* homolog was also found in Actinobacteria, which lack *mutS*/*mutL* homologs. A recent report showed that the *nucS* gene in *M. smegmatis* cells is required for mutation avoidance and antirecombination in the cells (17). This genetic study strongly supported the MMR function of EndoMS/NucS in this mycobacterial species. However, the purified EndoMS from *M. smegmatis* did not show the mismatch-specific endonuclease activity *in vitro*. Therefore, in the absence of combined biochemical and genetic studies in a single archaeal or actinobacterial species lacking canonical MutS/MutL-directed MMR proteins, the existence and specificity of an alternative MMR system was not fully understood or demonstrated.

Here, we have linked biochemical and genetic evidence for the mutation avoidance activities of the EndoMS family of proteins in *C. glutamicum*. We provide convincing evidence indicating that the physical interaction between the *C. glutamicum* replication β-clamp (DnaN) and the EndoMS homodimer is required for the recognition and cleavage of the mismatches *in vitro* and *in vivo* to protect corynebacterial cells from accumulation of mutations. Our biochemical data revealed that the physical interaction of the β-clamp substantially activates the mismatch-specific cleavage activity of the EndoMS/NucS proteins, and the carboxyl-terminal β-clamp-binding motif of *Cgl*EndoMS is necessary for this activation and mutation avoidance in the cells. *A priori*, the replication clamp could act directly on the catalytic step of EndoMS proteins or modulate its DNA binding properties. In the first scenario, the clamp could increase the catalytic rate constant up to four orders of magnitude, as has been demonstrated with other DNA repair nucleases (28). This is not fully compatible with our observations indicating 'only' four to five fold activation. In the future experiments, we will thus address whether the replication clamp helps EndoMS

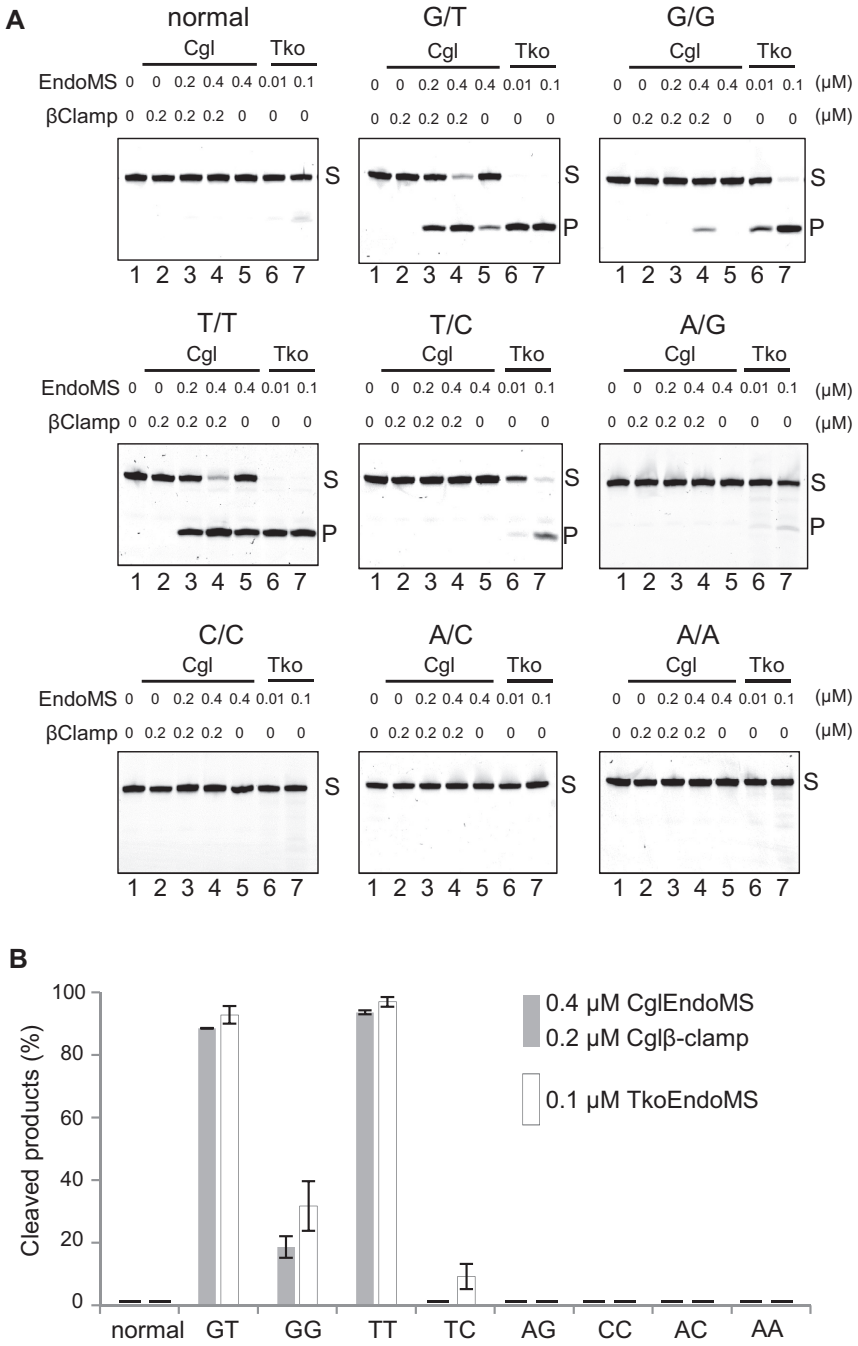

Figure 5. Preference for base-pair mismatches. (**A**) Five nanomolar Cy5-labeled dsDNA (45 bp), containing single base-pair mismatches (G/T, G/G, T/T, T/C, A/G, C/C, A/C, A/A), were incubated with various proteins as indicated on the panels. The products were separated by 10% PAGE. The representative results are shown. The base pairs are indicated on the top of each panel. The band assignments are indicated on the side of the panels, s, substrates; p, cleaved products. (**B**) Quantification of the cleaved products by 0.4 μM *Cgl*EndoMS and 0.2 μM *Cgl*β-clamp, and by 0.1 μM TkoEndoMS, as shown in lanes 4 and 7 in (A). The bars show the averages and the error bars are standard error of the mean (s.e.m.) from three independent experiments.

to scan DNA molecules that carry mismatches in order to recognize and process them.

Our observations already suggest a functional coupling of EndoMS/NucS-directed mismatch repair with the replisome. The role of the replication clamp, the β-clamp in Bacteria and PCNA in Archaea, is important for coordinating this non-canonical MMR, and thus diminishing the formation of unwanted double-strand chromosomal breaks by free EndoMS/NucS. After a double-strand break occurs at the mismatch site by EndoMS/NucS, this break would be processed by end resection to produce 3′-protruding ends, followed by homologous recombination. A recent articles showing that *Corynebacteria* are polyploidic organisms, resembling archaea (29), strongly supports the involvement of the homologous recombination in the repair of double

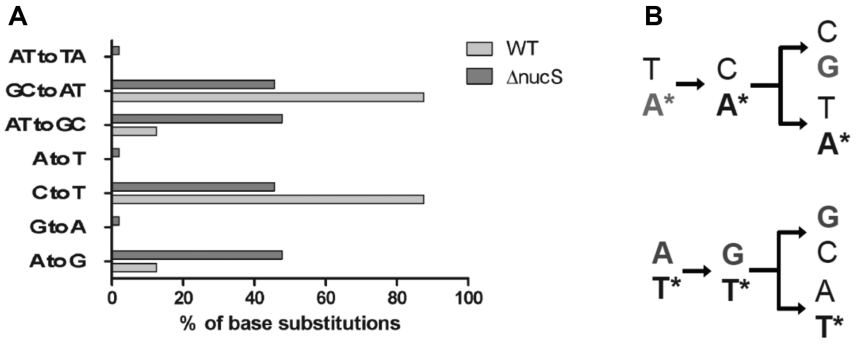

Figure 6. Mutational patterns of *C. glutamicum* wild type (WT) and Δ*nucS* strains. (**A**) Patterns of transition mutations (Rif^R) observed for wild type and Δ*nucS C. glutamicum* strains. (**B**) A simplified schema demonstrating how A to G transitions accumulating in the absence of *nucS* can result from either non-repaired G/T or C/A mismatches. Asterisks refer to the template strand bases.

strand breaks created around the repaired mismatches without strand discrimination.

*Cgl*EndoMS shows the order of preferences for G/T, T/T and G/G mismatches *in vitro* and these substrate preferences are quite similar to those of *Tko*EndoMS (11), signifying that the actinobacterial and archaeal EndoMS/NucS orthologues may share a common DNA repair function. The large majority of the Rif^R base substitutions occurring in the absence of EndoMS are transitions (Figure 6A) indicating that they must result from unrepaired purine mispairing with a wrong pyrimidine (G/T or C/A transition mismatches). Strikingly, whereas G/T mismatches act as preferred substrates for EndoMS/NucS proteins *in vitro* (Figure 5), C/A mismatches are processed very poorly (or not all). This is different from *E. coli* MutS/MutL system that corrects both transition mismatches [see for instance the early reference (30)]. The fact that both MMR systems work efficiently on the G/T mismatch is likely to have a biological meaning, as G/T mismatches are very common replication errors (8). This may reflect the relatively high capacity of at least some DNA polymerases to efficiently create G/T mismatches (31) that are stabilized by rare tautomeric forms of guanosine and thymidine (32). Transitions may thus occur more frequently in the cell than tranversions because molecular mechanisms generating and stabilizing them are different. It thus appears feasible that G/T mismatches frequently escaping proofreading activity of the replisome are efficiently repaired by the EndoMS:β-clamp complex. We stress that the EndoMS/NucS system is quite different from the very short patch repair system (Vsr) found for instance in *E. coli* (30). Although EndoMS and VSR endonucleases both efficiently cleave G/T mismatches, the Vsr system corrects G/T mismatches resulting from 5-methylcytosine deamination in a very specific sequence context. Moreover, the recognition of the G/T mismatch is achieved differently in the two systems. Cytosine methylation does not appear to occur in *Corynebacteria* and *Mycobacteria*, as experimentally suggested by PacBio sequencing (33,34). Our combined biochemical and genetic study thus suggests that EndoMS/NucS proteins have been fine-tuned to correct the frequently occurring transition mismatches in cells. We do not exclude the possibility that other types of mismatches (for instance T/T or G/G mismatches) resulting in transversions will also be processed by this novel DNA repair pathway, when they occur in cells.

In conclusion, we demonstrated the existence of a non-canonical MMR system in *C. glutamicum,* one of the most important industrial organisms used for the production of L-glutamate and L-lysine and other various products (35,36). The measured frequencies and the spectra of transition mutations for the wild type and Δ*nucS C. glutamicum* strains are consistent with the enzymatic properties of EndoMS/NucS-β–clamp complex *in vitro*. Our study may also be applicable for the industrial engineering of the useful strains and could provide a rational format for understanding the extent of genetic variation and the development of antibiotic resistant phenotypes in important human pathogens including *Mycobacterium tuberculosis*.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Takeshi Yamagami and Namiko Imai for technical assistance. We also thank Christine Houssin and Célia De Sousa-d'Auria (I2BC, Univ. Paris-Sud, France) for their expertise in *C. glutamicum* manipulation. We also thank Julia Gross for help with DNA sequencing experiments and Ursula Liebl for helpful comments on the manuscript.

FUNDING

Ministry of Education, Culture, Sports, Science and Technology of Japan [JP26242075 to Y.I.]; Japan Society for the Promotion of Science (JSPS) (to Y.I.); CNRS, INSERM and E. polytechnique (to H.M and S.S.). Funding for open access charge: INSERM.
Conflict of interest statement. None declared.

REFERENCES

1. Jackson,S.P. and Bartek,J. (2009) The DNA-damage response in human biology and disease. *Nature*, **461**, 1071–1078.
2. Ganai,R.A. and Johansson,E. (2016) DNA replication—a matter of fidelity. *Mol. Cell*, **62**, 745–755.
3. Jiricny,J. (2013) Postreplicative mismatch repair. *Cold Spring Harb. Perspect. Biol.*, **5**, a012633.

4. Tham,K.C., Kanaar,R. and Lebbink,J.H. (2016) Mismatch repair and homeologous recombination. *DNA Repair (Amst.)*, **38**, 75–83.
5. Matic,I., Rayssiguier,C. and Radman,M. (1995) Interspecies gene exchange in bacteria: the role of SOS and mismatch repair systems in evolution of species. *Cell*, **80**, 507–515.
6. Putnam,C.D. (2016) Evolution of the methyl directed mismatch repair system in Escherichia coli. *DNA Repair (Amst.)*, **38**, 32–41.
7. Pluciennik,A., Burdett,V., Lukianova,O., O'Donnell,M. and Modrich,P. (2009) Involvement of the beta clamp in methyl-directed mismatch repair in vitro. *J. Biol. Chem.*, **284**, 32782–32791.
8. Kunkel,T.A. and Erie,D.A. (2015) Eukaryotic mismatch repair in relation to DNA replication. *Annu. Rev. Genet.*, **49**, 291–313.
9. Pluciennik,A., Dzantiev,L., Iyer,R.R., Constantin,N., Kadyrov,F.A. and Modrich,P. (2010) PCNA function in the activation and strand direction of MutLalpha endonuclease in mismatch repair. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 16066–16071.
10. Lin,Z., Nei,M. and Ma,H. (2007) The origins and early evolution of DNA mismatch repair genes–multiple horizontal gene transfers and co-evolution. *Nucleic Acids Res.*, **35**, 7591–7603.
11. Ishino,S., Nishi,Y., Oda,S., Uemori,T., Sagara,T., Takatsu,N., Yamagami,T., Shirai,T. and Ishino,Y. (2016) Identification of a mismatch-specific endonuclease in hyperthermophilic Archaea. *Nucleic Acids Res.*, **44**, 2977–2986.
12. Nakae,S., Hijikata,A., Tsuji,T., Yonezawa,K., Kouyama,K.I., Mayanagi,K., Ishino,S., Ishino,Y. and Shirai,T. (2016) Structure of the EndoMS-DNA Complex as mismatch restriction endonuclease. *Structure*, **24**, 1960–1971.
13. Ariyoshi,M. and Morikawa,K. (2016) A dual base flipping mechanism for archaeal mismatch repair. *Structure*, **24**, 1859–1861.
14. Creze,C., Ligabue,A., Laurent,S., Lestini,R., Laptenok,S.P., Khun,J., Vos,M.H., Czjzek,M., Myllykallio,H. and Flament,D. (2012) Modulation of the Pyrococcus abyssi NucS endonuclease activity by replication clamp at functional and structural levels. *J. Biol. Chem.*, **287**, 15648–15660.
15. Meslet-Cladiere,L., Norais,C., Kuhn,J., Briffotaux,J., Sloostra,J.W., Ferrari,E., Hubscher,U., Flament,D. and Myllykallio,H. (2007) A novel proteomic approach identifies new interaction partners for proliferating cell nuclear antigen. *J. Mol. Biol.*, **372**, 1137–1148.
16. Ren,B., Kuhn,J., Meslet-Cladiere,L., Briffotaux,J., Norais,C., Lavigne,R., Flament,D., Ladenstein,R. and Myllykallio,H. (2009) Structure and function of a novel endonuclease acting on branched DNA substrates. *EMBO J.*, **28**, 2479–2489.
17. Castaneda-Garcia,A., Prieto,A.I., Rodriguez-Beltran,J., Alonso,N., Cantillon,D., Costas,C., Perez-Lago,L., Zegeye,E.D., Herranz,M., Plocinski,P. *et al.* (2017) A non-canonical mismatch repair pathway in prokaryotes. *Nat. Commun.*, **8**, 14246.
18. Portevin,D., De Sousa-D'Auria,C., Houssin,C., Grimaldi,C., Chami,M., Daffe,M. and Guilhot,C. (2004) A polyketide synthase catalyzes the last condensation step of mycolic acid biosynthesis in mycobacteria and related organisms. *Proc. Natl. Acad. Sci. U.S.A.*, **101**, 314–319.
19. Campbell,E.A., Korzheva,N., Mustaev,A., Murakami,K., Nair,S., Goldfarb,A. and Darst,S.A. (2001) Structural mechanism for rifampicin inhibition of bacterial rna polymerase. *Cell*, **104**, 901–912.
20. Ikeda,M. and Nakagawa,S. (2003) The Corynebacterium glutamicum genome: features and impacts on biotechnological processes. *Appl. Microbiol. Biotechnol.*, **62**, 99–109.
21. Sali,A. and Blundell,T.L. (1993) Comparative protein modelling by satisfaction of spatial restraints. *J. Mol. Biol.*, **234**, 779–815.
22. Laskowski,R.A., Rullmannn,J.A., MacArthur,M.W., Kaptein,R. and Thornton,J.M. (1996) AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. *J. Biomol. NMR*, **8**, 477–486.
23. Pettersen,E.F., Goddard,T.D., Huang,C.C., Couch,G.S., Greenblatt,D.M., Meng,E.C. and Ferrin,T.E. (2004) UCSF chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.*, **25**, 1605–1612.
24. Dalrymple,B.P., Kongsuwan,K., Wijffels,G., Dixon,N.E. and Jennings,P.A. (2001) A universal protein-protein interaction motif in the eubacterial DNA replication and repair systems. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 11627–11632.
25. Johanson,K.O. and McHenry,C.S. (1980) Purification and characterization of the beta subunit of the DNA polymerase III holoenzyme of Escherichia coli. *J. Biol. Chem.*, **255**, 10984–10990.
26. Garibyan,L., Huang,T., Kim,M., Wolff,E., Nguyen,A., Nguyen,T., Diep,A., Hu,K., Iverson,A., Yang,H. *et al.* (2003) Use of the rpoB gene to determine the specificity of base substitution mutations on the Escherichia coli chromosome. *DNA Repair (Amst.)*, **2**, 593–608.
27. Schaaper,R.M. and Dunn,R.L. (1987) Spectra of spontaneous mutations in Escherichia coli strains defective in mismatch correction: the nature of in vivo DNA replication errors. *Proc. Natl. Acad. Sci. U.S.A.*, **84**, 6220–6224.
28. Hutton,R.D., Roberts,J.A., Penedo,J.C. and White,M.F. (2008) PCNA stimulates catalysis by structure-specific nucleases using two distinct mechanisms: substrate targeting and catalytic step. *Nucleic Acids Res.*, **36**, 6720–6727.
29. Bohm,K., Meyer,F., Rhomberg,A., Kalinowski,J., Donovan,C. and Bramkamp,M. (2017) Novel chromosome organization pattern in Actinomycetales-Overlapping replication cycles combined with diploidy. *MBio*, **8**, e00511-17.
30. Kramer,B., Kramer,W. and Fritz,H.J. (1984) Different base/base mismatches are corrected with different efficiencies by the methyl-directed DNA mismatch-repair system of E. coli. *Cell*, **38**, 879–887.
31. Batra,V.K., Beard,W.A., Pedersen,L.C. and Wilson,S.H. (2016) Structures of DNA polymerase mispaired DNA termini transitioning to Pre-catalytic complexes support an Induced-Fit fidelity mechanism. *Structure*, **24**, 1863–1875.
32. Kimsey,I.J., Szymanski,E.S., Zahurancik,W.J., Shakya,A., Xue,Y., Chu,C.C., Sathyamoorthy,B., Suo,Z. and Al-Hashimi,H.M. (2018) Dynamic basis for dG*dT misincorporation via tautomerization and ionization. *Nature*, **554**, 195–201.
33. Blow,M.J., Clark,T.A., Daum,C.G., Deutschbauer,A.M., Fomenkov,A., Fries,R., Froula,J., Kang,D.D., Malmstrom,R.R., Morgan,R.D. *et al.* (2016) The epigenomic landscape of prokaryotes. *PLoS Genet.*, **12**, e1005854.
34. Zhu,L.X., Zhong,J., Jia,X.M., Liu,G., Kang,Y., Dong,M.X., Zhang,X.L., Li,Q., Yue,L.Y., Li,C.D. *et al.* (2016) Precision methylome characterization of Mycobacterium tuberculosis complex (MTBC) using PacBio single-molecule real-time (SMRT) technology. *Nucleic Acids Res.*, **44**, 730–743.
35. Heider,S.A. and Wendisch,V.F. (2015) Engineering microbial cell factories: Metabolic engineering of Corynebacterium glutamicum with a focus on non-natural products. *Biotechnol. J.*, **10**, 1170–1184.
36. Wendisch,V.F., Jorge,J.M.P., Perez-Garcia,F. and Sgobba,E. (2016) Updates on industrial production of amino acids using Corynebacterium glutamicum. *World J. Microbiol. Biotechnol.*, **32**, 105.
37. Jakoby,M., Ngouoto-Nkili,C.E. and Burkovski,A. (1999) Construction and application of new Corynebacterium glutamicum vectors. *Biotechnol. Tech.*, **13**, 437–441.
38. Menard,R., Sansonetti,P.J. and Parsot,C. (1993) Nonpolar mutagenesis of the ipa genes defines IpaB, IpaC, and IpaD as effectors of Shigella flexneri entry into epithelial cells. *J. Bacteriol.*, **175**, 5899–5906.