A type IV modification-dependent restriction enzyme SauUSI from Staphylococcus aureus subsp. aureus USA300

XU, Shuang-Yong, et al.

Abstract
A gene encoding a putative DNA helicase from Staphylococcus aureus USA300 was cloned and expressed in Escherichia coli. The protein was purified to over 90% purity by chromatography. The purified enzyme, SauUSI, predominantly cleaves modified DNA containing 5mC and 5-hydroxymethylcytosine. Cleavage of 5mC-modified plasmids indicated that the sites 5mCNGS (S = C or G) are preferentially digested. The endonuclease activity requires the presence of adenosine triphosphate (ATP) or dATP whereas the non-hydrolyzable γ-S-ATP does not support activity. SauUSI activity was inhibited by ethylenediaminetetraacetic acid. It is most active in Mg(++) buffers. No companion methylase gene was found near the SauUSI restriction gene. The absence of a cognate methylase and cleavage of modified DNA indicate that SauUSI belongs to type IV restriction endonucleases, a group that includes EcoK McrBC and Mrr. SauUSI belongs to a family of highly similar homologs found in other sequenced S. aureus, S. epidermidis and S. carnosus genomes. More distant SauUSI orthologs can be found in over 150 sequenced bacterial/archaea genomes. Finally, we […]

Reference
XU, Shuang-Yong, et al. A type IV modification-dependent restriction enzyme SauUSI from Staphylococcus aureus subsp. aureus USA300. Nucleic Acids Research, 2011, vol. 39, no. 13, p. 5597-610

DOI : 10.1093/nar/gkr098
PMID : 21421560

Available at:
http://archive-ouverte.unige.ch/unige:25418
Disclaimer: layout of this document may differ from the published version.
A type IV modification-dependent restriction enzyme SauUSI from *Staphylococcus aureus* subsp. *aureus* USA300

Shuang-yong Xu1,*, Anna R. Corvaglia2, Siu-Hong Chan1, Yu Zheng1 and Patrick Linder2

1New England Biolabs, Inc. 240 County Road, Ipswich, MA 01938, USA and 2Department of Microbiology and Molecular Medicine, Center of Excellence in Bacteriology of the University of Geneva, University of Geneva, Faculty of Medicine, 1, rue Michel Servet 1211 Geneve 4, Switzerland

Received December 2, 2010; Revised February 3, 2011; Accepted February 4, 2011

ABSTRACT

A gene encoding a putative DNA helicase from *Staphylococcus aureus* USA300 was cloned and expressed in *Escherichia coli*. The protein was purified to over 90% purity by chromatography. The purified enzyme, SauUSI, predominantly cleaves modified DNA containing 5mC and 5-hydroxymethylcytosine. Cleavage of 5mC-modified plasmids indicated that the sites S5mCNGS (S = C or G) are preferentially digested. The endonuclease activity requires the presence of adenosine triphosphate (ATP) or dATP whereas the non-hydrolyzable γ-S-ATP does not support activity. SauUSI activity was inhibited by ethylenediaminetetraacetic acid. It is most active in Mg++ buffers. No companion methylase gene was found near the SauUSI restriction gene. The absence of a cognate methylase and cleavage of modified DNA indicate that SauUSI belongs to type IV restriction endonucleases, a group that includes *EcoK* MrBC and Mrr. SauUSI belongs to a family of highly similar homologs found in other sequenced *S. aureus*, *S. epidermidis* and *S. carnosus* genomes. More distant SauUSI orthologs can be found in over 150 sequenced bacterial/archaea genomes. Finally, we demonstrated the biological function of the type IV REase in restricting 5mC-modified plasmid DNA by transformation into clinical *S. aureus* strain SA564, and in restricting phage λ infection when the endonuclease is expressed in *E. coli*.

INTRODUCTION

Restriction enzymes (REases) are grouped into four major types based on subunit architecture, adenosine triphosphate/guanosine triphosphate (ATP/GTP) requirement, sequence specificity and DNA cleavage mechanism (1). Type I restriction–modification (R–M) systems were originally discovered by studying phage plating efficiency among different *Escherichia coli* hosts (2). Type I restriction enzymes are multi-subunit complexes consisting of M2R2S subunits, and their activity requires ATP hydrolysis (3). Type II REases with 4–8-bp recognition sequences and over 300 unique specificities are widely used in creating recombinant DNA molecules (4). Type III R–M systems are also multi-subunit complexes with R2M2 configuration and require ATP hydrolysis for endonuclease activity (5–7). Two recognition sites in head-to-head or tail-to-tail orientations are essential for type I and type III enzymes to cleave. Modification-dependent REases are presently grouped into type IIM if the target sequence recognition and cleavage are very specific and precise, for example, DpnI, GN6mA (8) and GlaI, G5mC G5mC (9), MspJI, 5mCNNR (10); or type IV if cleavage is non-specific and variable. Examples of type IV are McrBC, which cleaves 5mC-modified DNA, and GmrSD that attacks glucosylated-hmC (glc-5hmC) T4 DNA (11–15).

Methicillin-resistant *Staphylococcus aureus* (MRSA) presents a great health threat by hospital-acquired infections (HaMRSA) and more recently by community acquired (CaMRSA) infections characterized by the spread of highly invasive and persistent infections (16,17). Recently, the ability of MRSA strains to acquire vancomycin-resistance genes from *Enterococcus* species

*To whom correspondence should be addressed. Tel: +1 978 380 7287; Fax: +1 978 921 1350; Email: xus@neb.com

© The Author(s) 2011. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/2.5), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
poses a significant threat (18–21). A large number of studies have been carried out to elucidate the mechanism of antibiotic resistance. The general conclusion is that drug resistance genes can be chromosomally encoded or derived from extra chromosomal elements, and acquired by horizontal gene transfer of mobile genetic elements such as conjugative plasmids, transposable elements, integrons or by phage infection and integration (22). The presence of type I, and II R–M systems or type IV restriction systems in clinical S. aureus strains could provide a genetic barrier to the free genetic exchange by transformation, transduction and conjugation (23,24).

A new gene that limits gene transfer was identified in two clinical S. aureus strains (SA564 and UAMS-1) (24). This open reading frame (ORF) contains a superfamily II DNA helicase domain and was found to be responsible for biological restriction of plasmid transformation when DNA was prepared from E. coli sources and suggested to be a type III restriction endonuclease based on the presence of helicase superfamily II domain. Inactivation of this gene by mutation increased the transformation efficiency by $10^2$–$10^4$-fold. However, no companion methylase of this gene by mutation increased the transformation efficiency, transduction and conjugation (23,24).

MATERIALS AND METHODS

Bacterial strains, genomic DNA and plasmids

ER2566 (C2566, NEB) is an E. coli B strain with the T7 RNA pol gene (T7 gene 1) integrated into the chromosome and deficient in Dcm methylase (Dcm), Mrr. The IMPACT protein expression kit and pTYB1 vector were obtained from NEB. Unmodified (Dcm-) λ phage DNA, pUC19 and pBC4 were prepared from ER2566. The backbone vector of pBC4 is pUC19 which carries a BstBI/ClaI fragment of adenovirus-2 DNA. T4gt DNA (deficient in x-and β-glucosyltransferases, containing 5hmC) was a gift from R. Vaisvila (NEB). λ Dcm’Dam’, λ Dcm’Dam’ and pBR322 Dam’ DNA’s were from NEB. Genomic DNA of S. aureus subsp. aureus USA300, FPR3757 was purchased from ATCC. Plasmid DNA was prepared by Qian mini-prep spin columns. Transformations of S. aureus cells with plasmid DNA were performed as described in Corvaglia et al. (24).

Protein expression and purification

The SauUSIR gene (2859 bp, 953 aa) flanked by SalI and XhoI restriction sites was amplified by polymerase chain reaction (PCR) using high-fidelity Phusion DNA polymerase (Finzymes). Following SalI and XhoI restriction digestion, the amplified DNA was ligated to pTYB1 with the same cohesive ends. The ligated DNA was transferred into ER2566 competent cells by transformation. The inserts in two clones with the highest enzyme activity were sequenced and found to contain the wild-type (wt) sequence. For small-scale protein purification, 4 L cells of ER2566 [pTYB1-SauUSIR] were cultured at 37°C in LB supplemented with Amp (100 µg/ml) until late log phase. isopropyl-β-D-thiogalactopyranoside (IPTG) induction was carried out at 0.5 mM final concentration and cell cultures induced at 16°C overnight. Cells were harvested by low-speed centrifugation, resuspended in a sonication buffer (20 mM Tris–HCl, pH 8.5, 0.5 M NaCl, 0.1% Triton X-100) and lysed by repeated sonication. After centrifugation, clarified cell lysate was loaded onto a chitin column (~21 ml, ~4 cm × 2.6 cm), which was then extensively washed with 10 column volumes of column buffer (20 mM Tris–HCl, pH 8.5, 0.5 M NaCl). The column was then flushed quickly with 20 ml of column buffer with 50 mM DTT and stored at 4°C overnight to allow the cleavage of SauUSIR from the SauUSIR–intein–chitin-binding domain (CBD) fusion. Approximately 15 ml of proteins (1 ml × 15) were eluted and analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE). To remove the nucleic acids from the eluted proteins, the pooled fractions were loaded onto a Source Q column (GE Healthcare) at 0.25 M NaCl concentration. The flow-through containing the SauUSIR protein was then loaded onto a HiTrap™ heparin column (GE Healthcare), and proteins were eluted with a linear gradient of 50 mM to 1 M NaCl. After analysis of the eluted fractions on SDS–PAGE, the fractions with SauUSIR were concentrated in a protein concentrator (Amicon/Millipore, 10000 kDa) by low-speed centrifugation. The protein was diluted in NEB Diluent A buffer and stored at −20°C.

To remove the divalent metal ions bound by the purified SauUSIR, the enzyme was supplemented with 20 mM EDTA and then dialyzed against a buffer (50 mM NaCl, 20 mM Tris–HCl, pH 8, 1 mM DTT) for 48 h at 4°C using a dialysis cassette (10000 Da MWCO; Thermo Scientific Pierce).
For batch purification of wt SauUSI and variants using chitin beads, 10 ml of cells were cultivated to late log phase at 37°C. After addition of 0.5 mM IPTG (final concentration), SauUSI protein expression was induced at 16°C overnight. Cell pellets were resuspended in 1 ml of cell lysis buffer (20 mM Tris–HCl, pH 8.5, 0.5 M NaCl, 0.1% Triton X-100), and cells were lysed by sonication. Clarified cell lysates were mixed with 200 μl of chitin beads prewashed with column buffer (20 mM Tris–HCl, pH 8.5, 0.5 M NaCl) and gently mixed on a rocker at 4°C for 1 h. Chitin beads with bound proteins were spun down at low speed centrifugation and washed three times with 1.5 ml of column buffer. The chitin beads were then resuspended in cleavage buffer (20 mM Tris–HCl, pH 8.5, 0.5 M NaCl, 50 mM DTT) and incubated at 4°C overnight. The supernatants containing SauUSI variants were analyzed on SDS-PAGE and assayed for cleavage activity on T4gt DNA.

For constitutive expression of SauUSI endonuclease gene (sauUSIR) in E. coli, a PCR fragment with re-engineered ribosome binding site (GGATCC-GGAGGT-AATGA-start codon) was cloned into the BamHI site of pBR322 under the TcR promoter. As a control, the sauUSIR gene inserted in the opposite orientation was also constructed.

Site-directed mutagenesis

The catalytic residues H119, K121, N139 and E150 of SauUSI endonuclease were mutated to alanine (Ala) by two-step overlapping PCR using Phusion™ high-fidelity enzymes. For medium-scale purification, cell lysates from 10 ml IPTG-induced overnight culture (16°C) were used for small-scale affinity purification, cell extracts from 10 ml IPTG-induced overnight cell cultures (16°C) were used for batch purification of the mutant enzymes. For medium-scale purification, cell lysates of 1 L IPTG-induced overnight cell cultures (16°C) were used for chitin column purification. Partially purified SauUSI variant H119A was used for gel filtration chromatography to determine its oligomerization state in NEB buffer 4.

SauUSI digestion and in vitro DNA methylation

SauUSI digested plasmid DNA was used as substrate for SauUSI digestion and cleavage site determination. SauUSI-digested PCR DNAs were purified by spin column and used as template for run-off sequencing to determine the cleavage sites.

Gel filtration study of SauUSI H119A variant

Ten microliters of purified SauUSI variant H119A (1 mg/ml) was injected into a Superdex 200 5/150 GL (3 ml bed volume; GE Healthcare). The column was run with NEB buffer 4 at 4°C at a flow rate of 0.2 ml/min. The partition coefficient (Kav) of SauUSI-H119A was determined by

\[
K_{av} = (v_v - v_o) / (v_t - v_v)
\]

where v is the elution volume, v is the void volume of the column, v is the total column volume. v and v were determined empirically by the addition of blue dextran and DTT, respectively to the sample. A standard curve was created by obtaining the Kav value of the standard proteins ovalbumin (43 kDa), conalbumin (75 kDa) and ferritin (440 kDa) (GE Healthcare) run under the same condition.

Phage plating and restriction of phage by SauUSI

Dcmλvir (a gift from A. Fomenkov, Mern Sibley, Lise Raleigh) was prepared from E. coli host ER1793. EcoK Dcm methylase modifies the DNA sequence CCWGG to produce modified sites C5mC (5 m-dCTP) or 5hmC (5hm-dCTP) instead of dCTP. Methylated plasmid DNAs were purified by spin column and used as substrate for SauUSI digestion. 5mC- or 5hmC-modified PCR DNAs with 5mCNGS sites (1, 2, 4, 5 and 6 sites) were produced in PCRs with 5mC (5 m-dCTP) or 5hmC (5hm-dCTP) instead of dCTP. The small amount of unmodified template DNA (˂0.1 ng/μl) did not interfere with SauUSI digestion and cleavage site determination. SauUSI-digested PCR DNAs were purified by spin columns using a PCR purification kit (Qiagen) and run-off sequencing was carried out using a Big-Dye terminator cycle sequencing kit (ABI/Life technologies). Similarly, after digestion of pBR-5hiHIM (G5mCNGC) plasmid by SauUSI endonuclease, the digested DNA was purified by spin column and used as template for run-off sequencing to determine the cleavage sites.

RESULTS AND DISCUSSION

SauUSI expression and purification

The sauUSIR gene was amplified from genomic DNA of S. aureus subsp. aureus USA300, FPR3757, and inserted into the T7 expression vector pTYB1. This results in a fusion of the ORF to the Sce VMA intein and Bacillus circulans CBD (IMPACT protein purification system, NEB). IPTG-induced ER2566 cell extracts showed endonuclease activity on λ DNA (Dcmλ) in the presence of ATP and 10 mM DTT (data not shown). Two clones with the correct-sized insert and the highest endonuclease activity were chosen for purification. The catalytic residues H119, K121, N139 and E150 of SauUSI endonuclease were mutated to alanine (Ala) by site-directed mutagenesis and H119A was used for gel filtration chromatography to determine its oligomerization state in NEB buffer 4.

SauUSI digestion and in vitro DNA methylation

SauUSI digested plasmid DNA was used as substrate for SauUSI digestion and cleavage site determination. SauUSI-digested PCR DNAs were purified by spin column and used as template for run-off sequencing to determine the cleavage sites.

Gel filtration study of SauUSI H119A variant

Ten microliters of purified SauUSI variant H119A (1 mg/ml) was injected into a Superdex 200 5/150 GL (3 ml bed volume; GE Healthcare). The column was run with NEB buffer 4 at 4°C at a flow rate of 0.2 ml/min. The partition coefficient (Kav) of SauUSI-H119A was determined by

\[
K_{av} = (v_v - v_o) / (v_t - v_v)
\]

where v is the elution volume, v is the void volume of the column, v is the total column volume. v and v were determined empirically by the addition of blue dextran and DTT, respectively to the sample. A standard curve was created by obtaining the Kav value of the standard proteins ovalbumin (43 kDa), conalbumin (75 kDa) and ferritin (440 kDa) (GE Healthcare) run under the same condition.

Phage plating and restriction of phage by SauUSI

Dcmλvir (a gift from A. Fomenkov, Mern Sibley, Lise Raleigh) was prepared from E. coli host ER1793. EcoK Dcm methylase modifies the DNA sequence CCWGG to produce modified sites C5mCNGC (W = A or T). Dcm−λvir was prepared from ER2566, a natural Dcm-deficient B strain. T4gt phage containing 5hmC was prepared from ER2566. Phages were diluted in phage broth or 10 mM MgCl2 solution and used to infect the E. coli hosts.

RESULTS AND DISCUSSION

SauUSI expression and purification

The sauUSIR gene was amplified from genomic DNA of S. aureus subsp. aureus USA300, FPR3757, and inserted into the T7 expression vector pTYB1. This results in a fusion of the ORF to the Sce VMA intein and Bacillus circulans CBD (IMPACT protein purification system, NEB). IPTG-induced ER2566 cell extracts showed endonuclease activity on λ DNA (Dcmλ) in the presence of ATP and 10 mM DTT (data not shown). Two clones with the correct-sized insert and the highest endonuclease activity were chosen for purification. The catalytic residues H119, K121, N139 and E150 of SauUSI endonuclease were mutated to alanine (Ala) by site-directed mutagenesis and H119A was used for gel filtration chromatography to determine its oligomerization state in NEB buffer 4.
activity were sequenced and found to carry the wt gene. A small-scale purification procedure was developed. SauUSI was first purified from a chitin column using the DTT-intein cleavage mechanism recommended by the manufacturer. The cleaved and eluted fractions were passed through a Source Q column at moderate salt concentration to remove most of the nucleic acids. The flow-through was then loaded onto a HiTrap™ heparin column and protein eluted with a NaCl gradient. The purification results are shown in Supplementary Figure S1. Two heparin fractions with the highest purity (Supplementary Figure S1, lanes 5 and 6) were concentrated to 0.2 mg/ml and used for further characterization.

**SauUSI substrate specificity**

Since SauUSI contains a helicase domain and ATP-binding site (inferred from amino acid sequence similarity to other type I and III REases), we included ATP (1 mM) in all digestions (see below for further data on nucleotide requirement). Figure 1A shows that SauUSI cleaves 5hmC-containing T4gt DNA as well as 5mC-containing DNA from a variety of sources: Dcm-modified λ and pBC4 DNA (pBC4 is a pUC19 derivative carrying a BstBI/ClaI fragment insert from adenovirus DNA), HeLa and Jurkat genomic DNA (human). *Escherichia coli* Dcm modifies CC(A/T)GG sequences at the internal cytosine, while mammalian DNA carries methylation in some CpG sites. Dam modification does not inhibit cleavage since both Dam+ and Dam− λ DNA were equally cleavable by SauUSI (Figure 1A, lanes 3 and 4). SauUSI displays lower cleavage activity on 5-glc-hmC DNA (glucosylated T4 DNA, wild type) (Figure 1A, lane 1). Unmodified (Dcm−) pBC4 is a poor substrate for SauUSI although a small fraction of DNA was partially linearized (Figure 1A, lane 5), which suggests that SauUSI has a non-specific endonuclease activity at high enzyme concentration, or that there is a low level of contaminating non-specific nuclease. The 0.4–0.5-kb DNA of T4gt digested products was an artifact which dispersed in high percentage agarose gel or PAGE (data not shown).

Figure 1B shows the undigested DNA substrates. SauUSI also displays very low endonuclease activity on Dam+ Dcm− pUC19 and pBR322 DNAs (data not shown). The specific activity of SauUSI is determined to be ~8000 U/mg protein on T4gt phage DNA in the presence of ATP. One unit of SauUSI is defined as the amount of enzyme required to digest 1 μg T4gt DNA into fragments of <500 bp in NEB buffer 4 at 37°C for 1 h.

*Figure 1.* SauUSI digestion of modified and unmodified DNA substrates. (A) DNA substrates digested by SauUSI endonuclease in 1X NEB buffer 4 at 37°C for 1 h. (B) Uncleaved DNA substrates. T4 wt DNA contains glucosylated 5-hmC DNA. T4gt deficient in α- and β-glucosyl transferases contains 5hmC DNA. Lane 9, 2-log DNA size marker. (C) SauUSI digestion of *in vitro* modified DNA. Plasmid pBC4 was methylated *in vitro* by the indicated methyltransferase and subsequently digested by SauUSI. Lane 7, mock-modified pBC4 DNA (Dcm−) (only SAM was present, no methylase). Lane 8, *in vivo* modified pBC4 (Dcm+). (D) *In vitro* methylated DNA (uncleaved) of the same substrates as in (C).
To further investigate SauUSI’s substrate specificity, Dcm\(^{-}\) plasmid pBC4 substrates were modified \textit{in vitro} by purified methylases and then used as substrates for SauUSI digestion. Figure 1C shows that SauUSI cleaves pBC4 plasmid after \textit{in vitro} modification by M.MspI (5mC(CCGG)), M.HpaII (G5mCCG), M.HhaI (G5mCGC) or M.Alul (A5mCT). Modified pBC4 DNAs by M.SsI (5mCG) or M.CviPI (G5mC) also allowed partial cleavage by SauUSI (data not shown). N4mC modified DNA by M.BamHI (GGAT\textsubscript{3}mCGG) or N6mA modified by M.TaqI (TCGA\textsubscript{6}mA) are poor substrates for SauUSI (Figure 1C, lanes 4 and 7). Dcm\(^{-}\) pBC4 (G5mCGWG site modified \textit{in vivo}) was again cleaved by SauUSI. In summary, SauUSI appears to cleave when 5mC is present in the following \textit{3} sequence contexts: 5mCC, 5mCG, 5mCT and 5mCW (see more sequence requirement below). SauUSI does not cleave N4mC and N6mA-modified DNA (the Dcm\(^{-}\) substrates do carry adenine modification in the G\textsubscript{6}mATC context). This 5mC/\textit{hmC}-dependent substrate specificity is shared by McrBC, although McrBC activity requires GTP hydrolysis. The digestion pattern, appearing as a ‘smeargram’, suggests non-specific nuclease activity (or strong ‘star’ activity) following incision from (or near) the modified 5mC/\textit{hmC} sites. The non-specific nuclease activity might be important to completely destroy the foreign invading DNA (the cleavage products are not easily repaired by simple religation).

We also tested SauUSI digestion of plasmids that have been modified \textit{in vivo} by 5mC methylases. Methylase-containing plasmids were co-transformed with pBC4 into a Dcm-deficient host ER2566 and plasmid mixture was prepared and digested with SauUSI in the presence of ATP. All plasmids were found to be mostly resistant to SauUSI (data not shown). N4mC modified pBC4 plasmid after SauUSI digestion. Figure 1C shows that SauUSI cleaves pBC4 substrates for SauUSI. In a control experiment, Dcm\(^{+}\) pBR322 or pBC4 additionally modified \textit{in vitro} by M.MspI (5mC(CCGG)) serves as a better substrate for SauUSI digestion than Dcm\(^{-}\) modified alone (data not shown). The recognition sequence of SauUSI is therefore somewhat similar to BisI (G5mC/NGC) and BslI (G5mCN/GC) (9). The enzymes are distinct, however: SauUSI cleaves DNA outside of its recognition sequence (see below) and requires ATP or dATP cofactor. Both enzymes require divalent cations, but ATP does not stimulate BisI activity (Priscilla Hiu-Mei Too and Shuang-yong Xu, unpublished data). In addition, the molecular mass of the purified BisI endonuclease was estimated to be \(\sim 21\, \text{kDa (\pm 2\, \text{kDa})}\) and the apparent molecular mass of SauUSI is \(\sim 100\, \text{kDa}\).

**ATP and dATP requirement for endonuclease activity**

Based on BlastP and amino acid sequence alignment (28) SauUSI contains a predicted ATP binding and DNA helicase domain in the middle part of the protein (approximately at 201–625 aa). To characterize the presumed nucleotide requirement, we tested SauUSI activity on T4gt DNA using different NTP or dNTP in the digestion. Figure 3A shows that only ATP or dATP supports SauUSI endonuclease activity. A non-hydrolyzable ATP analog, γ-S-ATP, does not support SauUSI endonuclease activity. The substrate specificity of the SauUSI endonuclease does not appear to cleave modified N4mC DNA efficiently.

**and bamHIM were poorly digested by SauUSI** (Figure 2C), indicating N4mC-modified DNAs are not good substrates for SauUSI. In a control experiment, Dcm\(^{+}\) plasmid pXbaI was cleaved by SauUSI.

The substrate specificity of the \textit{in vitro} and \textit{in vivo} modified DNA is listed in Table 1. The relative activity of SauUSI is summarized as follows: S5mCNGG (M.Fnu4HI, M.BssKI, M.Msp, M.Dcm) > G5mCTC (M.SacI), A5mGT (M.HpyCH4IV) >> 5mCAAT or 5mCATT (M.BsrDI). The best \textit{in vitro} DNA substrates are T4gt, pBR\textsubscript{3}m4HIM (G5mCNGG) and pACYC-bssKIM (G5mCNGG), which may reflect the frequency of the modified sites. For the limited number of N4mC-modified plasmids that have been tested, SauUSI endonuclease does not appear to cleave modified N4mC DNA efficiently.

The density of the 5mC-modified sites also appears to influence SauUSI endonuclease activity. Dcm\(^{+}\) pBR322 or pBC4 additionally modified \textit{in vitro} by M.MspI (5mC(CCGG)) serves as a better substrate for SauUSI digestion than Dcm\(^{-}\) modified alone (data not shown). The recognition sequence of SauUSI is therefore somewhat similar to BisI (G5mC/NGC) and BslI (G5mCN/GC) (9). The enzymes are distinct, however: SauUSI cleaves DNA outside of its recognition sequence (see below) and requires ATP or dATP cofactor. Both enzymes require divalent cations, but ATP does not stimulate BisI activity (Priscilla Hiu-MeiToo and Shuang-yong Xu, unpublished data). In addition, the molecular mass of the purified BisI endonuclease was estimated to be \(\sim 21\, \text{kDa (\pm 2\, \text{kDa})}\) and the apparent molecular mass of SauUSI is \(\sim 100\, \text{kDa}\).
Salt and metal ion requirement of SauUSI

The N-terminus of SauUSI contains a phopholipase D (PLD) family endonuclease catalytic site (the HKD catalytic motif, approximately at 1–200 aa) (30). Some PLD family nucleases are active in the absence of divalent metal ions (in the presence of EDTA) (31,32). We therefore evaluated whether SauUSI requires divalent metal ions for optimal activity. SauUSI is active in the four NEB buffers, all of which contain 10 mM Mg⁺⁺ (before EDTA treatment and dialysis), with highest activity in buffer 4 (50 mM potassium acetate). It is also active at NaCl concentrations of 50, 100, 150 or 200 mM with 10 mM MgCl₂ (Figure 4A, lanes 2 and 3, 5 and 6). SauUSI is partially active in a buffer in the absence of additional metal ions (50 mM NaCl, 10 mM Tris–HCl, pH 8, 1 mM DTT) as shown in Figure 4A, lane 8, but addition of 10 mM EDTA completely inhibited endonuclease activity (Figure 4A, lane 7). To clarify the nature of the requirement, we prepared enzyme depleted of EDTA-chelatable divalent cations. A volume of 20 mM EDTA was added into the SauUI enzyme, which was then dialyzed into a new buffer for 48 h. The metal ion-depleted SauUSI was retested for enzyme activity in buffer with or without divalent cations. Figure 4A, lane 9 shows that SauUSI is inactive after EDTA treatment and the subsequent extensive dialysis. The metal ion-depleted SauUSI regained activity in buffer 4 with Mg⁺⁺, and its activity was again inhibited if EDTA was added (Figure 4A, lanes 10 and 11). It is not clear whether the metal ion is required for the endonuclease activity or for the ATPase activity or for both. Further experiments are needed to uncouple the two reactions.

We also supplemented the depleted enzyme preparation with Mn⁺⁺, Ca⁺⁺, Co⁺⁺, Zn⁺⁺, Ni⁺⁺ or Cu⁺⁺, finding partial activity in buffers with 1 mM Mn⁺⁺, Ca⁺⁺ or Co⁺⁺. The relative endonuclease activity in different metal ions are Mg⁺⁺ > Mn⁺⁺ > Ca⁺⁺ > Co⁺⁺ (data not shown).
SauUSI is inactive in buffers with metal ions Zn$$^+$$, Ni$$^+$$ or Cu$$^+$$ (1–5 mM, data not shown). The promiscuous use of metal ions for SauUSI activity may provide certain advantage in restricting phage or plasmid DNA if so, activation might allow action in single-stranded DNA. However, this nuclease activity is not modification dependent (data not shown).

**Digestion of 5mC-containing PCR DNA substrates and run-off sequencing of the cleavage products**

A ~430-bp fragment of pACYC184 between the BamHI and SalI site that contains six SCNGS sites was generated by PCR amplification using 5m-dCTP or regular dCTP. The 5mC-containing PCR product was apparently digested by SauUSI endonuclease, generating a mixture of discrete bands and smearing (Figure 5A). The digestion was incomplete because many of the bands are bigger than the predicted cleavage products of 5mCCGG, 5mCGG, 5mCNCGG, 5mCNGG or 5mCWGG generated by virtual digestion using the NEB cutter (http://tools.neb.com/NEBcutter2/). In contrast, the PCR product made using regular dCTP was only much less cleaved at the same enzyme concentrations, confirming the preference of SauUSI to 5mC-modified DNA (Figure 5A).

Similarly, we generated a single site (GCCGC) 73-bp PCR fragment in PCR from pBR322 template and used it as a substrate for SauUSI digestion. The single-site PCR DNA was not cleaved by SauUSI (Figure 5B, lane 2). A two-site PCR substrate (GCTGC and GCAGG) with a length of 90 bp was cleaved by SauUSI and another two-site substrate (GCGGG, GCCGC) with 99 bp long was also sensitive to SauUSI digestion (Figure 5B, lanes 4 and 6). Longer 5mC-containing PCR products with 5

---

**Table 1. Digestion of 5mC-modified plasmid DNAs by SauUSI**

| C5 methylase | In vivo/in vitro modification | Modified sites | Relative endonuclease activity |
|-------------|-------------------------------|----------------|-------------------------------|
| M.HpaII     | *In vitro*                    | C Concert      | +                             |
| M.MspI      | *In vitro*                    | C Concert      | +                             |
| M.EagII     | *In vivo*                     | CG Concert     | +                             |
| M.BspKI     | *In vivo*                     | C Concert      | +                             |
| M.BsrFI     | *In vivo*                     | RC Concert     | +                             |
| M.BbvI      | *In vivo*                     | G Concert      | +                             |
| M.Fnu4HI    | *In vivo*                     | G Concert      | +                             |
| M.Dcm       | *In vivo*                     | C Concert      | +                             |
| M.TseI      | *In vivo*                     | G Concert      | +                             |

Consensus sequence: S ConcertNGS

| C5 methylase | In vivo | In vitro or in vitro | Modified sites | Relative endonuclease activity |
|-------------|---------|----------------------|----------------|-------------------------------|
| M.HhaI      | *In vitro* | G Concert | +                 |
| M.SssI      | *In vivo* | C Concert | +                 |
| M.AcI       | *In vivo* | AA Concert | +                 |
| M.HpyAV     | *In vivo* | A Concert | +                 |
| M.Aul       | *In vitro* | AG Concert | +                 |
| M.NgoHI     | *In vivo* | RC Concert | +                 |
| M.TspRI     | *In vivo* | C Concert | +                 |
| M.SacI      | *In vivo* | GAG Concert | +                 |
| M.BarDI     | *In vivo* | G Concert | +                 |
| M2.BarDI    | *In vivo* | G Concert | +                 |

(?) indicates possible modified cytosine; "+/-" indicates poor activity; "-" denotes no apparent activity in 1 h digestion, underlined C indicates modified base. R = A; G; S = G/C; W = A/T. The best-modified substrates are G5mCNGC (M.Fnu4HI) and 5mCNCGG (M.BspKI).
and 4 sites were also subjected to SauUSI digestion and partially cleaved products were detected (Figure 5B, lanes 8 and 10). Similar results were obtained with 5hmC-containing PCR DNAs. A single site (GCCGC) with 5hmC incorporated in PCR was a poor substrate for SauUSI digestion (Figure 5C), whereas 5hmC-containing PCR DNA with four SauUSI sites were cleavable by SauUSI (Figure 5D). 5hmC-containing PCR DNA with two sites was also cleaved by SauUSI (data not shown).

Run-off sequencing from the digested PCR products indicates that SauUSI cleaves DNA 2–3 bp outside of its recognition sequence GCCGC/GCGGC. There are three types of cleavage detected by run-off sequencing. Primer F210 detected two breaks (two double peaks) at the 3' side of GCGGC recognition site (Supplementary Figure S3, bottom panel) in comparison with undigested pBR322 sequence (top panel) (the cut site distance can be interpreted differently considering the overlapping GCCGC site). Primer F571 detected two breaks (two double peaks) from the 5' side of GCGGC/GCCGC target site (Supplementary Figure S4, bottom panel) compared to the control DNA sequence using the same primer. Primer F670 detected multiple DNA breaks/double peaks (Supplementary Figure S5), cleavage possibly taking place 3-bp 5' and 3' of the GCGGC/GCGGC target site, and another possible cut at 13-bp upstream (one DNA helical turn apart), implying that SauUSI is capable of sliding and cutting at long distance.

We also analyzed the cleavage sites of 5mC-modified plasmid DNA. The fnu4HIM gene was cloned in pBR322 and its expression was driven by the TcR promoter. The plasmid DNA pBR322-fnu4HIM extracted form overnight culture was fully resistant to Fnu4HI digestion (data not shown). Following digestion of pBR322-fnu4HIM by SauUSI for 2 h, the digested DNA was purified and subjected to run-off sequencing. Supplementary Figure S6 (bottom panel) shows that SauUSI cleaves outside of recognition sequence at N3, N10 and N14–18, with complete run-off at N18 from the target site, indicating SauUSI's ability to travel and cleave at long distance from the G5mCGC site. Supplementary Figure S6 (top panel) shows the DNA sequencing result from uncut pBR322 DNA. More cleavage sites remain to be analyzed in order to get a complete picture of SauUSI cut sites. The preliminary results indicate that SauUSI endonuclease is most likely
to cleave 2–3 bp outside of its recognition sequence on short PCR DNA (~90–200 bp). For the M.Fnu4HI-modified plasmid DNA (~6 kb), the SauUSI cleavage sites vary from N3 to N18 (G5mCNGC N3–18, bottom strand). It would be interesting to engineer SauUSI variants to cleave at longer distance at the fixed position.

Site-directed mutagenesis of catalytic residues of SauUSI endonuclease

The H119 and K121 residues in SauUSI are highly conserved among PLD family endonucleases; and the N139 and E150 residues are somewhat conserved (N, D, Q, E are sometimes interchangeable at the two positions). We mutated H119, K121, N139 and E150 to Ala by site-directed mutagenesis and partially purified the mutant proteins. Figure 6A shows the purified variant proteins.

The mutant protein yields are 4–5-fold higher than the wt enzyme, probably reflecting the reduced toxicity of the mutant endonucleases toward the E. coli expression host. The mutant enzyme variants are inactive in digestion of 5hmC T4gt DNA (Figure 6B, lanes 1–8). In contrast, the wt enzyme (batch-purified SauUSI in lanes 9 and 10, and column-purified SauUSI in lane 11) is active in cleaving T4gt DNA. This mutagenesis experiment confirmed the prediction that H119, K121, N139 and E150 are involved in catalytic activity. Mutations of these four catalytic residues to Ala abolished the endonuclease activity.

Oligomerization state of variant H119A

The oligomerization state of SauUSI was investigated using gel-filtration chromatography on an analytical
Superdex 200 column. The partition coefficient ($K_{av}$) of SauUSI H119A was compared to those of standard proteins to obtain the relative molecular weight ($M_r$). The elution profile of SauUSI H119A from two separate runs contains a single peak with $M_r$ of 76 kDa (open circles; Figure 7). The absence of peaks that correspond to oligomers of SauUSI indicates that SauUSI exists most likely as monomers in solution in the absence of target DNA. The physical properties of SauUSI enzyme with specific and non-specific DNAs remain to be analyzed.

**Restriction of phage by SauUSI endonuclease**

The biological function of type IV restriction systems is to restrict modified foreign DNA (phage or plasmid) (35,15). We tested the restriction of phages by SauUSI endonuclease using Dcm$^+$ or Dcm$^-$ phages that have been grown on modification proficient or deficient E. coli hosts. The plating efficiency of Dcm$^+$ phages is 0.03–0.04% (two independent experiments) on a SauUSI-containing host, compared to the high efficiency on host containing empty vector (Table 2). As expected, SauUSI does not restrict unmodified Dcm$^-$ phages. A frame-shift mutant of SauUSI (a single base deletion at nt position 85) abolished the ability to restrict Dcm-modified $\lambda$ phages. The deletion mutant (frame-shift mutant) was found during screening of sauUSIR gene insert in pBR322.

**In vivo DNA restriction activity of S. aureus clinical strain SA564 and its restriction-deficient counterpart**

Due to the possible existence of other active type I and type IV restriction systems in the S. aureus USA300 strain, we tested the possibility of this type IV (referred to as type III-like in Corvaglia et al., 2010) restriction system restricting methylated DNA in vivo, using S. aureus clinical strain SA564. We measured the transformation efficiencies of plasmid prepared from different E. coli strains. We used strains deficient in adenine methylation ($\text{dam}^-$), deficient in cytosine methylation ($\text{dcm}^-$), a double
mutant (dam\(^-\) dem\(^-\)), or wild-type (wt, dam\(^+\) dem\(^+\) DH5\(\alpha\)) for both methylation systems. To exclude problems in transformation efficiency other than the restriction of the plasmid DNA, we also prepared plasmid DNA from RN4220, the permissive laboratory \(S.\) \(aureus\) strain.

We have previously shown that plasmid DNA from RN4220 can efficiently be transformed into SA564, independent of the presence or absence of the two restriction systems. Mixtures of plasmids prepared from \(E.\) \(coli\) and \(S.\) \(aureus\) were transformed into the clinical strain SA564 and its restriction-deficient counterpart (SA564 \(hsdR\) \(::\) targetron/type I-deficient, type IV-like::targetron/deficient in SauUSI-like REase) (24). The results presented in Table 3 show that the transformation efficiency of \(Dcm^+\)-modified DNA is reduced by more than three-order of magnitude in wt SA564 strain. In contrast, plasmid DNA prepared from \(dem^-\) deficient strains efficiently transformed wild-type and mutant SA564 strains. In comparison, the absence of \(dam\) methylation did not influence the transformation efficiency. These results clearly show that this new restriction system recognizes C-methylated DNA and restricts only \(Dcm^-\)-modified DNA \textit{in vivo}.

**In vivo restriction (toxicity) by constitutive expression of SauUSI endonuclease in \(Dcm^+\) \(E.\) \(coli\)**

Initially, we tested plasmid DNA transformation of pTYB1-sau\(USIR\) into \(Dcm^-\) or \(Dcm^+\) \(E.\) \(coli\) strains. The transformation efficiency of pTYB1-sau\(USIR\) into a \(Dcm^-\) strain is only 5–10-fold lower than the empty vector pTYB1 (data not shown). We attributed this low level of restriction to the low expression level of SauUSI from pTYB1 in \(E.\) \(coli\) under non-inducing condition (or poor \textit{in vivo} cleavage of the fusion protein SauUSI-teinert-CBD to produce active SauUSI). To circumvent this problem, we recloned the sau\(USIR\) gene into pBR322 with an optimal ribosome binding site using a \(dem^-\) deficient host so that the sau\(USIR\) gene is constitutively expressed under the \(TcR\) promoter. Another clone with the sau\(USIR\) gene inserted in the opposite orientation was also

**Table 2. Phage restriction activity of SauUSI endonuclease**

| Cell     | \(Dcm^-\) λ PFU/ml | Average PFU/ml | Plating efficiency |
|----------|---------------------|----------------|-------------------|
| Exp. 1. ER2566 [pBR322] | \(5.8 \times 10^9\) | \(6.0 \times 10^8\) | 100% |
| Exp. 1. ER2566 [pBR-sauUSIR] | \(2.0 \times 10^9\) | \(2.4 \times 10^5\) | 0.04% |
| Exp. 1. ER2566 [pBR-sauUSIR**] | \(5.3 \times 10^9\) | \(5.6 \times 10^8\) | 93.3% |
| Exp. 2. ER2566 [pBR-sauUSIR] | \(3.2 \times 10^9\) | \(2.9 \times 10^8\) | 100% |
| Exp. 2. ER2566 [pBR-sauUSIR**] | \(1.0 \times 10^9\) | \(3.1 \times 10^8\) | 107% |
| Exp. 2. ER2566 [pBR-sauUSIR**] | \(2.9 \times 10^9\) | \(3.1 \times 10^8\) | 107% |
| Exp. 2. ER2566 [pBR-sauUSIR**] | \(3.2 \times 10^9\) | \(3.1 \times 10^8\) | 107% |

Plasmid pBR-sau\(USIR\)** carries a single base deletion at nt position 85, causing a frame-shift mutation in the SauUSI endonuclease.
constructed. Both plasmid clones were used to transform Dcm\(^{−}\) E. coli host NEB10\(\beta\). Table 4 shows the transformation result. The transformation efficiency of pBR-sauUSIR plasmid was reduced by more than four orders of magnitude compared to the empty vector pBR322 or pBR-sauUSIR\(^{*}\) (insert in reverse orientation). Although they are not isogenic strains, pBR-sauUSIR plasmid transforms two E. coli B strains deficient in Dcm methylase at high efficiency (data not shown). This experiment has some biological implication in the transfer SauUSI-like restriction genes among bacteria. The results suggest that as long as SauUSI expression level is tightly controlled, sauUSI-like genes could be acquired by 5mC-methylase carrying hosts.

**Functional domain organization in SauUSI protein**

BlastP analysis of SauUSI to known proteins in Genbank showed that the protein contains at least three functional domains (see Supplementary Figure S7) (36). The N-terminal domain (approximately from 1 to 200 aa) contains the catalytic residues of the phospholipase D (PLD) endonuclease family, specifically at H\(_{119}\)-K\(_{121}\)-N\(_{139}\)-E\(_{150}\). The HxK residues are highly conserved in PLD family endonucleases. N/E/D are also conserved, but can be substituted for each other in some circumstances. The middle part of the protein (approximately from 201 to 625 aa) carries the ATPase/helicase domain that is highly conserved among DEXD superfamily DNA/RNA helicases. There is a conserved putative Mg\(^{2+}\) binding site near amino acid 250, suggesting that Mg\(^{2+}\) may be required for ATP binding and hydrolysis. The C-terminal region (approximately from 685 to 953 aa) belongs to the conserved domain family DUF3427 found in bacteria and archaea with unknown function. We speculate that this region is the specificity domain that recognizes 5mC or 5hmC. This domain can be found in 255 proteins in the DUF3427 cluster (www.kegg.org). An interesting conserved motif prediction program found the conserved motif of Type III REases (from 221 to 373 aa in SauUSI, Supplementary Figure S7, bottom panel), explaining why this group of enzymes was originally annotated as putative type III REases containing a DNA/RNA helicase domain in Genbank. We also noted that in some sequenced bacterial genomes, the PLD family endonuclease exists as a separate protein and a DNA/RNA helicase domain in an adjacent protein, possibly regulated by a single operon (data not shown). Site-directed mutagenesis of the conserved residues in the helicase/ATPase and 5mC/5hmC domains will need to be undertaken to confirm these predictions.

Interestingly, there is a conserved gene (SauUSA300_2432) linked to the SauUSI gene (SauUSA300_2431) that has strong amino acid sequence similarity to the pyrophosphohydrolase and MutT/NUDIX family hydrolases. This protein may be involved in the breakdown of pyrophosphate generated by ATP/dATP hydrolysis. The SSDB gene cluster search results from the KEGG server indicated that there are at least 60 SauUSI-like type IV endonucleases linked to the putative pyrophosphohydrolases (data not shown). The function of the latter protein remains to be examined.

### Table 4. Transformation efficiency of pBR-sauUSIR, pBR-sauUSIR\(^{*}\) and pBR322 plasmids into Dcm\(^{−}\) E. coli strain NEB10\(\beta\)

| Plasmid clone          | Comments | CFU/\(\mu\)g DNA | Average CFU/\(\mu\)g |
|------------------------|----------|------------------|----------------------|
| pBR322                 | Empty vector | 5.0 \(\times\) 10\(^6\) | 4.5 \(\times\) 10\(^6\) |
| pBR-sauUSIR\(^{*}\)    | R gene inserted in opposite orientation to Tc\(^{\alpha}\) promoter | 4.6 \(\times\) 10\(^6\) | 4.1 \(\times\) 10\(^6\) |
| pBR-sauUSIR            | R gene inserted under Tc\(^{\alpha}\) promoter | 4.5 \(\times\) 10\(^2\) | 3.3 \(\times\) 10\(^2\) |
| pBR-sauUSIR\(^{*}\)    | SauUSI frame-shift mutant | 4.4 \(\times\) 10\(^{4}\) | 3.3 \(\times\) 10\(^{4}\) |

Plasmid pBR-sauUSIR\(^{*}\) carries a single base deletion at nt position 85, causing a frame-shift mutation in the SauUSI endonuclease (the mutation was introduced in PCR).

### Table 3. Transformation efficiency (number of colony forming units) using dcm\(^{−}\) modified plasmid and dam\(^{−}\)/dcm\(^{−}\) DNA

| Exp.1 | Source of plasmid | Recipient strains | Source of plasmid |
|-------|-------------------|-------------------|-------------------|
|       | pCN33 DH5\(\gamma\)/pCN38 RN4220 | pCN33 BL21 dcm\(^{-}\)/pCN38 RN4220 | pCN33 LC1015 dam\(^{-}\)/pCN38 RN4220 | pCN33 LC3142 dam\(^{-}\), dcm\(^{-}\)/pCN38 RN4220 |
|       | 564 type IV\(^{-}\) | 1/71 100         | 28 500/99 100     | 20/45 500     | 10 200/57 900 |
|       | 564 type IV\(^{-}\) | 8900/30 100      | 8900/53 900       | 7800/32 100  | 7800/40 700  |

| Exp.2 | Source of plasmid | Recipient strains | Source of plasmid |
|-------|-------------------|-------------------|-------------------|
|       | pCN38 DH5\(\gamma\)/pCN33 RN4220 | pCN38 BL21 dcm\(^{-}\)/pCN33 RN4220 | pCN38 LC1015 dam\(^{-}\)/pCN38 RN4220 | pCN38 LC3142 dam\(^{-}\), dcm\(^{-}\)/pCN33 RN4220 |
|       | 564 type IV\(^{-}\) | 12/17 500        | 78 800/30 900     | 12/15 890   | 23 900/37 600 |
|       | 564 type IV\(^{-}\) | 34 100/7320      | 44 200/10 600     | 64 700/13 300 | 24 900/10 500 |

Note: Here, we refer to the SauUSI isoschizomer in the S. aureus clinical strain SA564 as a type IV restriction enzyme.
Inspection of sequenced *Staphylococcus* sp. genomes revealed a few type II R-M systems that may create potential substrates for SauUSI following 5mC modification. *Staphylococcus aureus* RF122 strain contains a putative type II R-M system, whose C5 methylase has significant amino acid sequence similarity to GCNGC methylases (62% aa sequence identity to M.Bsp6I, GCNGC). *Staphylococcus aureus* ST398 strain carries a putative R–M system with the predicated target sequence of CCNGG (the putative C5 methylase displays 69% aa sequence identity to M1.ScrFI, CCNGG). *Staphylococcus aureus* TW20 and *S. epidermidis* RP62A strains carry a putative orphan C5 methylase (possibly prophage encoded multi-specificity C5 methylases) with predicted GGCC specificity. These 5mC-modified DNA at GCNGC, CCNGG and GGCC sites can potentially be cleaved by SauUSI, and 5mC-modified DNA from these strains may be restricted when transferred into *S. aureus* USA300 strain or *Staphylococcus* species with active SauUSI-like REases. The susceptibility of these methylase-modified DNAs to SauUSI restriction remains to be tested.

SauUSI displays low activity on short duplex oligos and requires at least two sites for efficient cleavage, possibly SauUSI requiring extensive DNA looping and communication of two distantly bound molecules for optimal activity, as it was observed for type I and type III REases. In SauUSI, the C-terminal region is presumed to contain the 5mC/5hmC recognition domain. Work is underway to clone and express this TRD domain and to characterize its activity. It is possible to select SauUSI mutants that prefer to cleave 5mC DNA. Such mutants would be useful for epigenetic studies.

**CONCLUSIONS**

SauUSI is a 5mC/5hmC modification-dependent REase. The endonuclease activity requires ATP or dATP and divalent cations such as Mg++, Mn++, Ca++ or Co++. Although the cleavage appears to be non-specific, the initial incision (nicks or ds breaks) displays certain sequence preferences, with 5mCNGS as the preferred sites. Efficient cleavage requires more than one site; and the cleavage sites outside of its recognition sequence appear to be variable (N2–N18). SauUSI is shown to restrict Dcm+ phage infection and Dcm+ plasmid transformation in *E. coli* and *S. aureus* clinical strain SA564. SauUSI endonuclease expression is toxic when it is constitutively expressed in a Dcm− *E. coli* strain. SauUSI is a multi-domain protein with a PLD-family endonuclease catalytic site located at the N-terminal domain, a DNA helicase/ATPase located in the middle part of the protein, and a presumed 5mC/hmC TRD domain (DUF3427) located at the C-terminus. Over 150 proteins with 25–100% amino acid sequence identity with a similar functional domain organization are found in sequenced bacterial genomes.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online.

**ACKNOWLEDGEMENTS**

We thank Richard Roberts, Elisabeth Raleigh, and William Jack for their critical comments, and Rick Morgan for discussions. We are grateful to Lise Raleigh, Mern Sibley, Alexey Fomenkov, Rick Morgan, Keith Lunnen, Geoff Wilson, Lucia Greenough, Bill Jack for strains and plasmids. We are appreciative to NEB DNA sequencing lab for sequencing the SauUSI expression clones and SauUSI cleavage products. We express our gratitude to Don Comb and Jim Ellard for their continued support of basic research.

**FUNDING**

New England Biolabs, Inc. The Swiss National Science Foundation and the Canton of Geneva. Funding for open access charge: New England Biolabs, Inc.

Conflict of interest statement. None declared.

**REFERENCES**

1. Roberts, R.J., Belfort, M., Bestor, T., Bhagwat, A.S., Bickle, T.A., Bitinaite, J., Blumenthal, R.M., Degtyarev, S., Dryden, D.T., Dyvig, K. et al. (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.*, 31, 1805–1812.
2. Arber, W. (1974) DNA modification and restriction (Review). *Prog. Nucleic Acid Res. Mol. Biol.*, 14, 1–37.
3. Wilson, G.G. and Murray, N.E. (1991) Restriction and modification systems. *Annu. Rev. Genet.*, 25, 585–627.
4. Roberts, R.J., Vincze, T., Posfai, J. and Macelis, D. (2010) REBASE – a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.*, 38, D234–D236.
5. Meisel, A., Mackeldanz, P., Bickle, T.A., Kruger, D.H. and Schroeder, C. (1995) Type III restriction endonucleases translocate DNA in a reaction driven by recognition site-specific ATP hydrolysis. *EMBO J.*, 14, 2958–2966.
6. Humbelin, M., Suri, B., Rao, D.N., Hornby, D.P., Eberle, H., Pripfl, T., Kenel, S. and Bickle, T.A. (1988) Type III DNA restriction and modification systems EcoP1 and EcoP15. Nucleotide sequence of the EcoP1 operon, the EcoP15 mod gene and some EcoP1 mod mutants. *J. Mol. Biol.*, 200, 23–29.
7. Szczelkun, M.D., Friedhoff, P. and Seidel, R. (2010) Maintaining a sense of direction during long-range communication on DNA. *Biochem. Soc. Trans.*, 38, 404–409.
8. de la Campa, A.G., Springhorn, S.S., Kale, P. and Lackey, S.A. (1988) Proteins encoded by the DpnI restriction gene cassette. *J. Biol. Chem.*, 263, 14696–14702.
9. Tarasova, G.V., Nayakshina, T.N. and Degtyarev, S.K. (2008) Substrate specificity of new methyl-directed DNA endonuclease Gial. *BMC Mol. Biol.*, 9, 7.
10. Zheng, Y., Colen-Karni, D., Xu, D., Chin, H.G., Wilson, G., Pradhan, S. and Roberts, R.J. (2010) A unique family of Mrr-like modification-dependent restriction endonucleases. *Nucleic Acids Res.*, 38, 5527–5534.
11. Sutherland, E., Coe, L. and Raleigh, E.A. (1992) MerBC: a multisubunit GTP-dependent restriction endonuclease. *J. Mol. Biol.*, 225, 327–358.
12. Raleigh, E.A., Murn, N.E., Revel, H., Blumenthal, R.M., Westaway, D., Reith, A.D., Rigby, P.W.J., Elhai, J. and Hanahan, D. (1988) MerA and MerB restriction phenotypes of some *E. coli* strains and implications for gene cloning. *Nucleic Acids Res.*, 15, 1563–1575.
13. Mulligan, E.A., Hatchwell, E., McCormick, S.R. and Dunn, J.J. (2011) Differential binding of *Escherichia coli* MerA protein to DNA sequences that contain the dinucleotide m5CpG. *Nucleic Acids Res.*, 38, 1997–2005.
14. Mulligan, E.A. and Dunn, J.J. (2008) Cloning, purification and initial characterization of E. coli McrA, a putative 5-methylcytosine-specific nuclease. *Meth. Enzymol.*, 462, 98–103.

15. Bair, C.L. and Black, L.W. (2007) A type IV modification dependent restriction nuclease that targets glucosylated hydroxymethyl cytosine modified DNAs. *J. Mol. Biol.*, 366, 768–776.

16. Loomba, P.S., Taneja, J. and Mishra, B. (2010) Methicillin and vancomycin resistant *Staphylococcus aureus* in hospitalized patients. *J. Glob. Infect. Dis.*, 2, 275–283.

17. Deleo, F.R., Otto, M., Kreiswirth, B.N. and Chambers, H.F. (2010) A type III-like restriction endonuclease function, ToxIN, functions as a protein–RNA toxin–antitoxin pair. *Proc. Natl Acad. Sci. USA*, 106, 894–899.

18. Liu, G., Ou, H.Y., Wang, T., Li, L., Tan, H., Zhou, X., Rajakumar, K., Deng, Z. and He, X. (2010) Cleavage of phosphorothioated DNA and methylated DNA by the type IV restriction endonuclease ScoMcrA. *PLoS Genet.*, 6, e1001253.

19. Zhu, W., Clark, N.C., McDougal, L.K., Hageman, J., McDonald, L.C. and Patel, J.B. (2008) Vancomycin-resistant *Staphylococcus aureus* isolates associated with Inc18-like vanA plasmids in Michigan. *Antimicrob. Agents Chemother.*, 52, 452–457.

20. Weigel, L.M., Donlan, R.M., Shin, D.H., Jensen, B., Clark, N.C., McDougal, L.K., Zhu, W., Musser, K.A., Thompson, J., Kohlerschmidt, D. et al. (2007) High-level vancomycin-resistant *Staphylococcus aureus* isolates associated with a polymicrobial biofilm. *Antimicrob. Agents Chemother.*, 51, 231–238.

21. Mulligan, E.A. and Dunn, J.J. (2008) Cloning, purification and initial characterization of E. coli McrA, a putative 5-methylcytosine-specific nuclease. *Meth. Enzymol.*, 462, 98–103.

22. Fluit, A.C. and Schmitz, F.J. (2004) Resistance integrons and resistance integrons and potential active site residues. *Protein Sci.*, 5, 914–922.

23. Lunn, K. and Xu, S.Y. (2008) Expression and purification of BmrI restriction endonuclease and its N-terminal cleavage domain variants. *Protein Expr. Purif.*, 58, 42–52.

24. Lagunavicius, A., Sasauskas, G., Halford, S.E. and Siksnys, V. (2003) The metal-independent type II restriction enzyme BII is a dimer that binds two DNA sites but has only one catalytic centre. *J. Mol. Biol.*, 326, 1051–1064.

25. Bidner, E., Chopin, A., Ehrlich, S.D. and Chopin, M.C. (2009) Activation of mRNA translation by phage protein and low temperature: the case of *Lactococcus lactis* abortive infection system AbiD1. *BMC Mol. Biol.*, 10, 4.

26. Fineran, P.C., Blower, T.R., Foulds, I.J., Humphreys, D.P., Lilley, K.S. and Salmon, G.P. (2009) The phage abortive infection system, ToxIN, functions as a protein–RNA toxin–antitoxin pair. *Proc. Natl Acad. Sci. USA*, 106, 894–899.

27. Liu, G., Ou, H.Y., Wang, T., Li, L., Tan, H., Zhou, X., Rajakumar, K., Deng, Z. and He, X. (2010) Cleavage of phosphorothioated DNA and methylated DNA by the type IV restriction endonuclease ScoMcrA. *PLoS Genet.*, 6, e1001253.