Journal of Bioscience and Bioengineering
VOL. 127 No. 5, 529–538, 2019

www.elsevier.com/locate/jbiosc

High copy number mutants derived from *Corynebacterium glutamicum* cryptic plasmid pAM330 and copy number control

Shuhei Hashiro,[1] Mayu Mitsuhashi,[1] and Hisashi Yasueda[1,2,*]

Institute for Innovation, Ajinomoto Co., Inc., 1-1 Suzuki-cho, Kawasaki-ku, Kawasaki 210-8681, Japan[1] *and Research and Development Center for Precision Medicine, University of Tsukuba, 1-2 Kasuga, Tsukuba-shi, Ibaraki 305-8550, Japan*[2]

Received 19 July 2018; accepted 11 October 2018
Available online 9 November 2018

A high copy number mutant plasmid, designated pVC7H1, was isolated from an *Escherichia coli*–*Corynebacterium glutamicum* shuttle vector pVC7N derived from cryptic plasmid pAM330 that was originally found in *Brevibacterium lactofermentum* 2256 (formally *C. glutamicum* ATCC 13869). The copy number of pVC7N was estimated to be about 11 per chromosome, whereas pVC7H1 displayed a copy number of 112 per chromosome in *C. glutamicum*. The mutation (designated *copA1*) was in a region between long inverted repeats (designated the *copA1* region) and was identified as a single base conversion of cytosine to adenine. By introduction of a cytosine to guanine mutation (designated *copA2*) at the same site as *copA1*, a further high copy number mutant (>300 copies of the plasmid per chromosome) was generated. Through genetic and RNA-Seq analyses of the *copA1* region, it was determined that a small RNA (designated sRNA1) is produced from the upstream region of *repA*, a gene encoding a possible replication initiator protein, and sRNA1 is a possible regulator of the copy number of pAM330-replicon-contaning plasmids. Determination of the precise transcription start sites of sRNA1 and *repA*-mRNA suggested that sRNA1 could sequester a presumed ribosome binding site of *repA*-mRNA from ribosomes by an antisense RNA-mediated mechanism. Our data also indicate that the secondary-structure of sRNA1 is crucial for its function in plasmid copy number control.

[**Key words:** Antisense RNA-mediated mechanism; *Corynebacterium glutamicum*; High copy number mutant plasmid; pAM330; pBL1-family; Plasmid copy number control; RepA protein; Small RNA; U-turn motif]

Brevibacterium lactofermentum, a coryneform bacterium, has been used industrially for the fermentative production of various amino acids including L-glutamate (1,2). The strain was later classified as *Corynebacterium glutamicum* (3). Cryptic plasmid pAM330 was originally found in *B. lactofermentum* 2256 (ATCC 13869) in 1984. This plasmid consists of 4448 bp and contains eight putative open reading frames (ORFs), among which ORF-1 and ORF-5 are reported to be essential in its autonomous replication (4,5). Studies show that pAM330 is very similar (or even identical) to pBL1 independently isolated from *B. lactofermentum* BL0 (6). Thus, pAM330 is classified as a member of the corynebacterial plasmid pBL1-family, which replicates by a rolling circle replication (RCR) mechanism (7,8). ORF-1 encodes a replication initiator protein (Rep protein) and the gene is designated *rep*. However, ORF-5 has no homology with other known protein sequences and its function has not been elucidated (5,9). The copy number of pAM330 in *B. lactofermentum* was estimated to be between 10 and 30 per cell (5). Subsequently, for genetic engineering studies in *C. glutamicum*, a useful shuttle vector between *C. glutamicum* and *Escherichia coli* designated pVC7 (10) was constructed from pAM330 and pHSG399 (11) and can confer chloramphenicol tolerance (Cm^R) on both host strains. The copy number of pVC7 in *C. glutamicum* was comparable to that of pAM330.

Another type of cryptic plasmid from *C. glutamicum* ATCC 13058 is pHM1519 (4,9), a representative of the plasmid pCG1-family which includes pCG1 (8), a small cryptic plasmid found in strain *C. glutamicum* ATCC 31808 (12). Although pHM1519 replicates by the RCR mechanism, a pHM1519-replicon-containing plasmid is compatible with a pAM330-replicon-contaig plasmid within a *C. glutamicum* host cell (i.e., both can be used simultaneously). In a similar way to the construction of pVC7, another type of shuttle vector between *C. glutamicum* and *E. coli* was constructed from pHM1519 and pHSG299 (11) carrying a kanamycin resistance determinant (*kan*); the resultant plasmid was named pPK4 (13). While the copy number of pHM1519 in *C. glutamicum* was estimated to be 140 per cell by a classical method (4), pPK4 displayed a slightly higher copy number than that of pVC7-derived plasmid in *C. glutamicum* (14).

The replication control mechanism of small cryptic plasmids in coryneform bacteria has been intensively studied using pCG1 and pGA1. Plasmid pGA1 is also an RCR-type plasmid, isolated from *C. glutamicum* strain LP-6 (15); it belongs to the pCG1-family and is present at about 35 copies per chromosome in *C. glutamicum* cells. The minimal replicon of pGA1 has the *rep* gene coding for the replication initiator. In the upstream region of the *rep* gene, a small counter-transcribed RNA (ctRNA) was identified to be responsible for negative control of the plasmid copy number (16,17). Recently in pCGR2 and pCG1, members of the same family of plasmids, it was revealed that short antisense RNA molecules are encoded in the

* Corresponding author at: Institute for Innovation, Ajinomoto Co., Inc., 1-1 Suzuki-cho, Kawasaki-ku, Kawasaki 210-8681, Japan. Tel.: +81 44 245 8554; fax: +81 44 244 9617.
E-mail addresses: shuhei_hashiro@ajinomoto.com (S. Hashiro), mayu_mitsuhashi@ajinomoto.com (M. Mitsuhashi), hisashi_yasueda@ajinomoto.com (H. Yasueda).

1389-1723/$ – see front matter
https://doi.org/10.1016/j.jbiosc.2018.10.012

DNA strand complementary to the 5′-leader sequence of their respective *repA* genes (18). It was predicted that the short RNAs and their target *repA*-mRNAs could form highly structured molecules having stem-loops with U-turn motifs (19) which mediate fast RNA−RNA pairing, and thus it was indicated that the characteristic terminal loop structures in both RNAs are important for efficient antisense RNA−mRNA interaction by which the plasmid copy number is regulated within the host cells.

A variety of plasmid vectors will be needed for fundamental metabolic analysis in *C. glutamicum* and for construction of strains for useful material production in biotechnological applications. In particular, it is generally useful to express a target gene at a high gene dosage for the production of the gene-product in a large quantity, thus there is value in developing high copy number vectors. Therefore, we conducted this study on basic analysis of copy number control in the pAM330-replicon. At first, we resequenced both plasmids pAM330 and pVC7 stored in our laboratory and found that both have some changes in the DNA sequence from the original reported sequences. We give the names pAM330N and pVC7N to the newly identified plasmids, respectively, to distinguish them from the original pAM330 and pVC7. Subsequently, we obtained high copy number (*cop*) mutant plasmids derived from pVC7N in *C. glutamicum* and analyzed the mutations by genetic and molecular methods to elucidate the copy number control mechanism of this pAM330-replicon-containing plasmid. To the best of our knowledge, this is the first report of a high copy number mutant and elucidation of the possible copy number control mechanism in the *C. glutamicum* pBL1-plasmid family.

MATERIALS AND METHODS

Strains, plasmids, and media *C. glutamicum* 2256L is a strain derived from *C. glutamicum* 2256 by curing pAM330 (4); it was constructed using plasmid pVC7-sacB (described in the next section) and Dex-S10 agar selection medium as described previously (20,21). *C. glutamicum* strains were routinely grown on CM-Dex medium (20) at 30°C. *E. coli* JM109 was used as host for construction and preparation of recombinant plasmids and was grown in Luria−Bertani (LB) medium at 37°C with shaking. When appropriate, antibiotics were added as follows: Cm at 5 mg/L and 20 mg/L for *C. glutamicum* and *E. coli*, respectively; kanamycin (Km) at 25 mg/L for both strains. The plasmids used are listed in Table 1.

Plasmid construction All DNA primers used in this study were obtained from Eurofins Genomics (Tokyo, Japan) and are listed in Table S1. For construction of pVC7N derivatives (Table 1), inverse PCR was applied to introduce mutations and to delete target DNA regions, and also mutant plasmids containing base-substitutions or deletions were generated using a KOD-Plus Mutagenesis Kit (Toyobo, Osaka, Japan). The combinations of DNA primers and templates used for construction of pVC7N derivatives are listed in Table S2. Construction of pPK4 derivatives containing four fragments of the *copA1* region from pVC7N was performed by insertion into pPK4 (13) of DNA fragments amplified from pVC7N. A vector fragment of pPK4 was amplified using KOD FX Neo (Toyobo) and primers P68 and P69. DNA fragments (named F1-R2, F1-R3, F2-R1, and F3-R1) were amplified by PCR using primer pairs P70 and P74, P70 and P75, P71 and P73, and P72 and P73, respectively with pVC7N as the template. Then plasmids pPK4-F1R2, pPK4-F1R3, pPK4-F2R1, and pPK4-F3R1 were constructed by ligating the amplified DNA fragments using the In-Fusion HD Cloning Kit. Plasmid pVC7-sacB containing the *Bacillus subtilis sacB* gene as a counterselectable marker was constructed as follows. A DNA fragment containing *sacB* was amplified by PCR using primers P20 and P21 and pBS4S (22) as the template. A vector fragment was prepared by PCR using primers P22 and P23 and pVC7N as the template. Then the *sacB*-fragment and the vector fragment were ligated to construct pVC7-sacB.

Plasmid pGEM-RD2 was constructed from pGEM-T Easy vector (Promega, WI, USA), a *repA* gene fragment (named RepAf; 1056 bp), and a *dnaA* gene fragment amplified from *C. glutamicum* 2256L chromosomal DNA (named DnaAf; 794 bp).

TABLE 1. Plasmids used in this study.

Plasmid	Relevant characteristics	Source or reference
pAM330	Cryptic plasmid in *C. glutamicum* 2256 (ATCC 13869)	4,5
pAM330N	Cryptic plasmid in *C. glutamicum* 2256 stored in laboratory	This study
pVC7N	Shuttle vector derived from pAM330N and pHSG329, Cm^r	This study
pVC7NdUR	pVC7N derivative deleting DNA region (in nt positions 1 to 884) of pVC7N, Cm^r	This study
pVC7NdDR	pVC7N-derivative deleting DNA region (in nt positions 3245 to 4447) of pVC7N, Cm^r	This study
pVC7H1	pVC7N derivative, *copA1*, Cm^r	This study
pVC7H2	pVC7N derivative, *copA2*, Cm^r	This study
pVC7H3	pVC7N derivative, *copA3*, Cm^r	This study
pVC7H4	pVC7N derivative, *copA4*, Cm^r	This study
pVC7H5	pVC7N derivative, *copA5*, Cm^r	This study
pVC7H6	pVC7N derivative, *copA6*, Cm^r	This study
pVC7H7	pVC7N derivative, *copA7*, Cm^r	This study
pVC7-C1063d	pVC7N derivative, *mut22*, Cm^r	This study
pVC7-C1063G	pVC7N derivative, *mut21*, Cm^r	This study
pVC7-C1094d	pVC7N derivative, *mut20*, Cm^r	This study
pVC7-C1094G	pVC7N derivative, *mut19*, Cm^r	This study
pVC7-A1131d	pVC7N derivative, *mut18*, Cm^r	This study
pVC7-A1131T	pVC7N derivative, *mut17*, Cm^r	This study
pVC7-C1147d	pVC7N derivative, *mut16*, Cm^r	This study
pVC7-C1147G	pVC7N derivative, *mut15*, Cm^r	This study
pVC7-T1169d	pVC7N derivative, *mut14*, Cmr	This study
pVC7-T1169A	pVC7N derivative, *mut13*, Cm^r	This study
pVC7-C1183d	pVC7N derivative, *mut12*, Cm^r	This study
pVC7-C1183G	pVC7N derivative, *mut11*, Cm^r	This study
pVC7-1163d1183d	pVC7N derivative, *mut26*, Cm^r	This study
pVC7-1163C1183d	pVC7N derivative, *mut25*, Cm^r	This study
pVC7-1163d1183G	pVC7N derivative, *mut24*, Cm^r	This study
pVC7-1163C1183G	pVC7N derivative, *mut23*, Cm^r	This study
pHM1519	Cryptic plasmid in *C. glutamicum* ATCC 13058	4
pPK4	Shuttle vector derived from pHM1519 and pHSG229, Km^r	13
pPK4-F1R3	pPK4 containing copA region (nt 977 to 1206) of pVC7N, Km^r	This study
pPK4-F1R2	pPK4 containing copA region (nt 977 to 1252) of pVC7N, Km^r	This study
pPK4-F2R1	pPK4 containing copA region (nt 1034 to 1320) of pVC7N, Km^r	This study
pPK4-F3R1	pPK4 containing copA region (nt 1091 to 1320) of pVC7N, Km^r	This study
pBS4S	pHSG299 derivative containing *sacB* gene from *Bacillus subtilis*, Km^r	21
pVC7-sacB	pVC7 derivative containing *sacB* gene from pBS4S, Cm^r	This study
pGEM-T Easy	Cloning vector, Ap^r	Promega
pGEM-RD2	pGEM-T Easy derivative containing *repA* gene fragment from pAM330N and *dnaA* gene fragment from *C. glutamicum* 2256L	This study

Specifically, using pGEM-T Easy vector as the template and primers P84 and P85, a vector fragment was prepared using PrimeSTAR GXL DNA polymerase (Takara Bio, Shiga, Japan). RepAf was obtained by PCR amplification using primer pair P80 and P81 with pVC7N as template, and DnaAf was generated using P82 and P83 and KOD FX NEO (Toyobo). Fragments RepAf and DnaAf were concatenated and cloned into the pGEM-T Easy vector fragment using the In-Fusion HD Cloning Kit. *E. coli* JM109 was transformed with the reaction mixture, and pGEM-RD2 was obtained.

DNA isolation, manipulation, and transformation Plasmid DNA from *C. glutamicum* was isolated by a modified alkaline extraction procedure with lysozyme using the QIAprep Miniprep Kit (Qiagen, Santa Clarita, CA, USA). To isolate total DNA (chromosomal and plasmid DNA) from *C. glutamicum* harboring plasmid(s), cells were grown in CM-Dex medium at 30°C until middle-to-late exponential growth phase and cells of culture were harvested. The cells were resuspended in TE buffer (10 mM Tris−HCl, 1 mM EDTA, pH 8.0) containing lysozyme (Sigma−Aldrich, MO, USA), followed by incubation at 37°C for 30 min. Then, 2% SDS solution were added, and the mixture was subsequently treated with proteinase K (Takara Bio) at 56°C for 30 min. After heat-inactivation of proteinase K, RNaseA (Qiagen) was added and the mixture was incubated at room temperature for 5 min. Next, the mixture was treated with phenol-chloroform-isoamyl alcohol (25:24:1) (Nacalai Tesque, Kyoto, Japan) solution to extract DNA. The DNA was precipitated with 2-propanol after addition of sodium acetate, and then the DNA sample was rinsed with 70% ethanol. Finally, the precipitate was dissolved in TE buffer and the concentration of nucleic acids was measured using a NanoDrop 2000 spectrophotometer (Thermo Fisher Scientific Inc., MA, USA). Restriction enzymes, T4 DNA ligase, and the *E. coli* JM109 electroporation kit were purchased from Takara Bio, and used as recommended by the manufacturers. *C. glutamicum* was transformed by electroporation (23).

Determination of plasmid copy number Copy numbers of plasmids were determined using quantitative PCR (qPCR) and total DNA extract of *C. glutamicum* 2256L cells harboring each plasmid to be investigated; a relative quantification method was used, similar to that described by Tsuchida et al. (24). Two genes, *repA* in the objective pAM330-replicon-containing plasmid and *dnaA* on the *C. glutamicum* chromosome, were selected as targets for qPCR analysis, and plasmid copy number was calculated as the ratio of quantified *repA* to *dnaA*. Three pairs of DNA primers were used for probing each gene: for *repA*, P34 and P35, P36 and P37, and P38 and P39; for *dnaA*, P28 and P29, P30 and P31, and P32 and P33 (Table. S1). DNA fragments of RepAf and DnaAf in pGEM-RD2 were designed to contain all target regions used for the qPCR amplification. qPCR was carried out using a LightCycler 96 real-time PCR system (Roche Molecular Systems, Inc., CA, USA) with SYBR Premix EX *Taq* (Tli RNaseH Plus) (Takara Bio) according to the manufacturer's instructions. A tenfold serial dilution series of pGEM-RD2 was prepared ranging from 1×10^{-14} to 1×10^{-10} g/reaction tube for qPCR of *dnaA* and *repA* and standard curves displaying the cycle threshold parameter (Ct) values plotted against the log of the initial DNA concentration were generated. The relative copy numbers of *dnaA* and *repA* genes in samples were determined from qPCR-analysis using extracted total DNA with reference to the standard curves and mean values were calculated from values obtained for the three target sites for each gene. The plasmid copy number was compared to the chromosomal copy number.

Evaluation of plasmid stability Evaluation of plasmid stability was performed as described previously (14). In brief, the frequency of loss of the plasmid vectors after growth for about 60 generations in the absence or presence of selective antibiotic pressure was examined. *C. glutamicum* 2256L cells harboring plasmid were grown in CM-Dex medium, and cultures were diluted after every 12 h of growth and plated onto CM-Dex medium-agar plates. Then, the antibiotic resistance of at least 100 colonies grown on each plate was examined. Plasmid stability was expressed as: the number of colonies growing on chloramphenicol-containing plates/the total number of colonies on the non-selective plates. The structural intactness of plasmids was analyzed by agarose gel electrophoresis of plasmid DNA extracted from cells cultured for about 60 generations.

DNA sequencing analysis DNA sequencing analysis was performed on a Genetic Analyzer 3500 × L (Applied Biosystems, Foster City, CA, USA) with a BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems). Plasmids pAM330N and pVC7N were sequenced using DNA primers P01 to P19 (Table S1).

RNA-seq analysis *C. glutamicum* 2256L/pVC7N was grown in CM-Dex medium containing chloramphenicol at 30°C until late-logarithmic phase and the culture broth was mixed with RNAprotect Bacteria Reagent (Qiagen). After the cells were collected by centrifugation, the pellet was suspended in bacterial lysis buffer (10 mM Tris−HCl, pH 7.0, 1 mM EDTA, 15 g/L lysozyme), and the suspension was incubated at room temperature (RT) for 20 min. Then, proteinase K was added to the suspension at RT for a further 20 min. Total RNA isolation was performed using TRIzol LS Reagent (Thermo Fisher) according to the manufacturer's instructions and the RNA concentration was measured with a NanoDrop 2000. To remove residual DNA from the sample, the RNA extract was treated with an RNase-free DNase set (Qiagen), and the RNA fraction was again purified with TRIzol LS reagent. rRNA was removed from the RNA fraction using a Ribo-Zero rRNA Removal Gram-positive Kit (Illumina, Tokyo, Japan). Then, sequencing libraries were constructed with the TruSeq Stranded mRNA Sample Preparation Kit (Illumina) and sequencing was performed by Illumina MiSeq. FASTQ files as sequence data were imported into CLC Genomics Workbench software version 7.0.4 (Qiagen) as paired-end reads and trimmed based on quality using default settings. Then trimmed reads were aligned to the DNA sequence of pVC7N.

Rapid amplification of cDNA ends (5′-RACE) For mapping of the 5′-end of sRNA1, *C. glutamicum* 2256L/pVC7N was cultured and total RNA was prepared from the cells with TRIzol treatment. The RNA was polyadenylated using a poly(A) Tailing Kit (Thermo Fisher) and used in cDNA synthesis with 5′-RACE. The reaction was performed using the SMARTer RACE cDNA Amplification Kit and protocol (Takara Bio). In brief, first-strand cDNA was reverse transcribed from the poly(A)-tailed RNA using SMARTScribe Reverse Transcriptase (RTase) with the 5′-RACE cDNA synthesis primer and the SMARTer IIA oligo from the kit. An aliquot of the cDNA was then amplified using P76 (Table S1) as a forward target-specific primer and a universal primer mix from the kit. PCR conditions used were as described in the kit protocol. Then, nested PCR was performed using the universal primer and a nested target-specific primer, P77. The PCR product was purified using a MiniElute PCR Purification Kit (Qiagen) and the DNA sequence was analyzed.

For identification of the 5′-end of *rep*-mRNA, total RNA was prepared from *C. glutamicum* 2256L/pVC7H2. First-strand cDNA was synthesized using a *repA* gene-specific primer, P78, by RTase at 45°C. Then, using the cDNA as the template, SeqAmp PCR was performed using *repA* gene-specific primer P79 and universal primer mix provided in the kit, and the PCR product was purified and sequenced.

***In silico* analysis of gene detection and prediction of RNA secondary structure** The EasyGene server (http://www.cbs.dtu.dk/services/EasyGene/) (25) and GeneMarkS program (http://exon.gatech.edu/Genemark/genemarks.cgi) (26) were employed for gene detection. The secondary structure of RNA was predicted using the mfold web server (http://unafold.rna.albany.edu/?q=mfold/RNA-Folding-Form) (27) and the CentroidFold program (http://www.ncrna.org/centroidfold) (28).

Nucleotide sequence accession number The nucleotide sequence of pAM330N and pVC7N analyzed in this study were deposited into the DDBJ and GenBank databases with accession number LC406945 and LC425431, respectively.

RESULTS

Structural analysis of pAM330-replicon *C. glutamicum* plasmid pAM330N used in this study had alterations at 18 places in the DNA sequence compared to the original reported sequence of pAM330 (5), including deletions and insertions (Fig. S1). The total nt number of pAM330N was 4447 bp, and the sequence was identical to that of pBL1 (9) except for three consecutive bases in positions 3751 to 3753 in the pAM330N nucleotide sequence (Fig. S2). The sequence analysis revealed the presence of seven ORFs (named ORF-1 to ORF-7) with initiation codons ATG, GTG or TTG and also containing at least 300 nucleotides (Fig. 1). Thus, the organization of ORFs in pAM330N was also different from that in the original pAM330. A stored shuttle vector pVC7 constructed from pAM330 and pHSG399 was also resequenced to confirm its DNA sequence. The DNA sequence of the region derived from pAM330 within pVC7 completely matched that of pAM330N except for a single base substitution (C to T at position 1140, Fig. S1). Curiously the base T at position 1140 in the sequence of the stored pVC7 was identical to the corresponding base of pAM330, not pAM330N. Thus, we named this shuttle plasmid pVC7N, which was used for further experiments in this study.

Subsequently, to localize the autonomous replication region derived from pAM330N in *C. glutamicum*, we constructed two kinds of pVC7N-derived plasmids, namely pVC7dUR and pVC7dDR, in which we respectively deleted an upstream region (designated UR) and a downstream region (designated DR) of the ORF-1 region of pVC7N (Fig. 1). Each constructed shuttle plasmid was introduced into *C. glutamicum* strain 2256L. Transformants with pVC7dDR could be efficiently obtained and the copy number per cell was comparable to that of pVC7N. However, we could not transform *C. glutamicum* 2256L with pVC7dUR. This result indicated that the downstream region (DR) is not crucial for replication by the pAM330-replicon. Therefore, within the autonomous replication region of pAM330N, only two ORFs (ORF-1 and ORF-2) were found,

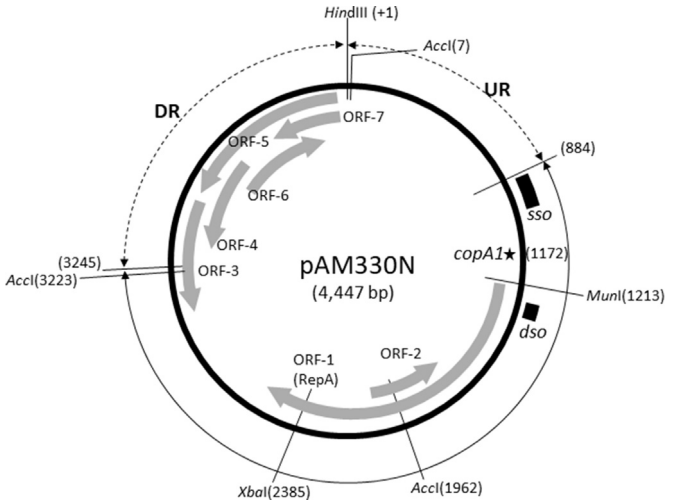

FIG. 1. Schematic structure of pAM330N. Predicted coding regions (ORF-1 to ORF-7) are marked by gray arrows indicating the direction. The positions of the single-stranded origin (*sso*) and the double-stranded origin (*dso*) are indicated in accordance with functional annotation in pBL1 (7). UR and DR indicate the regions deleted in the construction of pVC7N-dUR and pVC7N-dDR, respectively. The first nucleotide "A" in the *Hin*dIII recognition sequence (5′-AAGCTT-3′) is assigned as position +1, and representative restriction enzyme cleavage sites are shown in parentheses. The site of mutation *copA1* (position 1172) is marked with an asterisk.

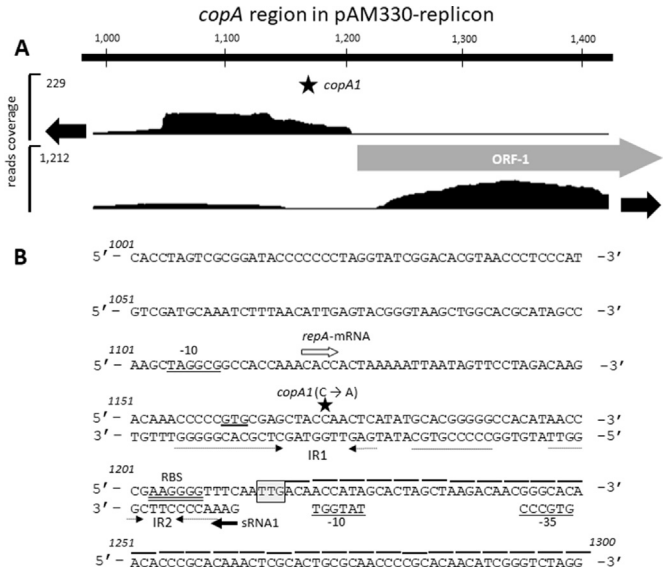

FIG. 2. *copA* region of pAM330-replicon. (A) RNA-Seq analysis of the *copA* region. Strand-specific RNA sequencing over the DNA region from position 980 to 1420 of pVC7N is shown. Upper and lower x-axes indicate antisense and sense strands in ORF-1 (indicated with a gray arrow), and the y-axis shows reads per nucleotide of RNA-Seq coverage. The position of *copA1* is indicated with an asterisk. (B) DNA segment from positions 1001 to 1300 of pVC7N. The transcriptional start points of *repA*-mRNA and sRNA1 are indicated with open and closed thick arrows, respectively, and the presumed promoter sequences (−35 and −10 regions) are shown. The inverted repeats IR1 and IR2 are marked with dotted arrows. The presumed ribosome binding site (RBS; 5′-AAGGGG-3′) for RepA is double-underlined and the predicted start codon TTG is boxed. The following reading frame for RepA is indicated with overlines. The triplet GTG at positions 1161 to 1163 is also underlined. The site of the *copA1* mutation (position 1172) is indicated with an asterisk.

although the short ORF-2 overlaps with ORF-1 in the opposite direction (Fig. 1). Although pAM330N was almost identical to pBL1 (9), we named ORF-1 *repA* (encoding the RepA replication initiator protein for the pAM330-replicon) to distinguish it from the gene named *rep* in pBL1.

In ORF-1, the longest possible ORF starts at GTG at positions 1161–1163, and the second and the third longest ORFs initiate at TTG at positions 1215–1217 (Fig. 2B) and ATG at positions 1395–1397, respectively. It was reported that ORF-1 encoded by the *rep* gene in the pBL1 plasmid sequence (GenBank: AAD08690.1) is composed of 401 amino acid residues starting from an ATG codon, and it corresponds to the ATG for ORF-1 in pAM330N. However, in this study, we deduced that the TTG codon is the initiation codon for RepA, rather than the GTG or ATG codons, for the following reasons: First, a single base deletion mutation at position 1169 or 1172, which would result in a frame-shift of the reading frame from the GTG codon at positions 1161–1163, did not significantly impair plasmid replication (this result is also shown in the following section), indicating that RepA is not encoded in this region. Second, a possible ribosome binding site (RBS) 5′-AAGGGG-3′ detected closely upstream of the TTG triplet is more similar to the consensus RBS (5'-(A)AGGAG-3′) (30) than a possible RBS (5′-AGGCA-3′) in front of the triplet ATG site. Third, in the upstream region of the TTG codon in pAM330N, the nucleotide sequence and structural features are very similar to those from another cryptic plasmid (pCC1) in *Corynebacterium callunae* (31) (Fig. 3). As shown in Fig. 3, the sequence, the location of the RBS for Rep and RepA, and also two stem-loop (SL) structures observed in the upstream region of *rep*-mRNA and *repA*-mRNA are common features in the genetic organization of pCC1 and pAM330N. Finally, ORF prediction by software program, EasyGene (25), trained on the gene database of *C. glutamicum* as the target microorganism, indicated that ORF-1 starts from the TTG codon. Thus, the translational initiation codon of RepA of the pAM330-replicon was presumed to be the TTG codon in positions 1215–1217 (Fig. 2B). TTG is often used as a start codon in *C. glutamicum* (29).

ORF-2, which could in principle encode a polypeptide of 147 amino acid residues, is located within ORF-1 but encoded by the reverse strand DNA (Fig. 1). The presumed initiation codon was also TTG, but no RBS-like sequence was found upstream of the TTG codon. When the putative polypeptide encoded by ORF-2 was compared with protein databases by BLAST searches (National Center for Biotechnology Information, http://www.ncbi.nlm.gov), no significant homology was found. In addition, the EasyGene program did not predict ORF-2 as a real gene. Thus, we tentatively concluded that ORF-2 does not encode a functional polypeptide. In addition, a small putative ORF (designated ORF-5 in former study (5)) capable of encoding 65 amino acid residues had been predicted to be essential for plasmid replication in previous work (5). However, in the present work, it was found that this small ORF is included within the long reading frame of ORF-1. EasyGene could not detect ORF-5 as a real gene. Therefore, we decided to exclude this small putative ORF from consideration as an essential factor for plasmid replication.

Isolation of *cop* mutant of pVC7N in *C. glutamicum* For investigation of replication control in the pAM330-replicon, we tried to isolate high copy number (*cop*) mutants of pVC7N. Initially, the copy numbers of pAM330N and pVC7N in *C. glutamicum* were determined to be about 20 and 11 per chromosome, respectively. Subsequently, we screened *cop* mutants of pVC7N formed by spontaneous mutation. Since pAM330N is stably maintained in *C. glutamicum* 2256, pVC7N was also introduced into the cell and the transformants were screened on agar medium containing 10 mg/L chloramphenicol (higher than the concentration commonly used in selection of *C. glutamicum* transformants). We isolated 60 transformants forming relatively large colonies on this agar medium. Among them, one candidate pVC7N mutant with obviously increased copy number was obtained by examining the plasmid content in

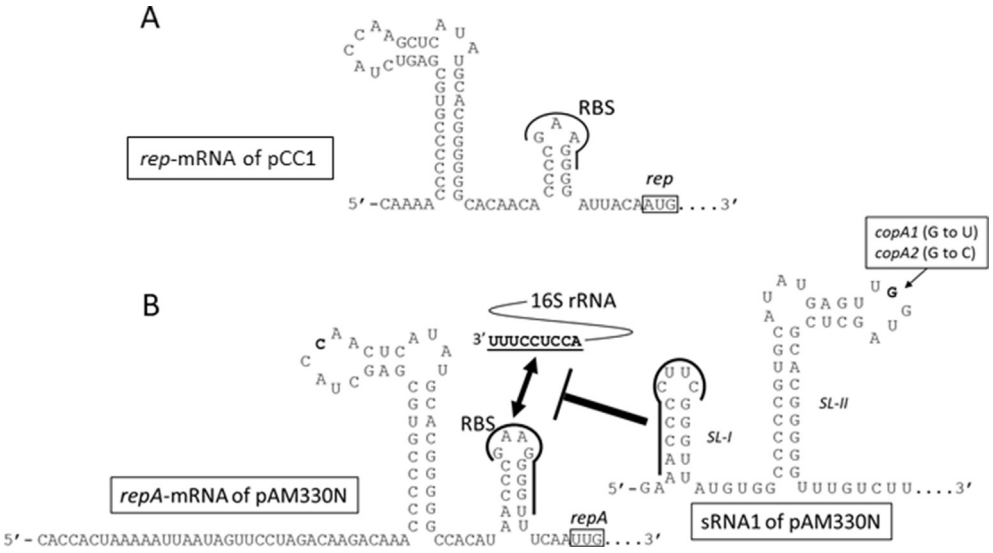

FIG. 3. Model of copy number control in pAM330-replicon plasmid. (A) A predicted *rep*-mRNA leader region in pCC1 (31) which belongs in the same pBL1 plasmid family as pAM330N is shown. The possible RBS and initiator codon (AUG) are indicated. Interestingly, this region resembles the *copA* region of pAM330N, although the putative initiation codon of pCC1-Rep is AUG not UUG. (B) *repA*-mRNA leader sequence in pAM330-replicon plasmid, pVC7N, shown with presumed RBS and initiator codon UUG for RepA. Translation of *repA*-mRNA starts through the interaction between the putative RBS and the 3′-region (3′-UUUCCUCCA-5′) of 16S rRNA. A sequence 5′-ACCCCUUC-3′ complementary to the putative RBS of *repA*-mRNA is located in SL-I at the 5′-end of sRNA1, which could interfere with the translational initiation of *repA*-mRNA by RNA−RNA interaction. The position of *copA1* and *copA2* is also indicated in the sRNA1 sequence.

each clone. Most other candidates forming colonies on the selection medium plate were transformants that lost pAM330N and maintained only pVC7N. The high copy number candidate plasmid was introduced into a new host cell of *C. glutamicum* 2256L and we confirmed that the plasmid was a *cop* mutant; the copy number was estimated to be 112 ± 6 per chromosome. This mutant plasmid was named pVC7H1 (Fig. 4A). Subsequently, we identified the *cop* mutation of pVC7H1 in the region (designated the *copA1* region) between a long inverted-repeat known as a characteristic hairpin structure (5) (Fig. 2B) located upstream of ORF-1 (*repA*) and found that the mutation was transversion of cytosine to adenine at position 1172 of the pVC7N DNA sequence (Figs. 1, 2 and 4B). This *cop* mutation was named *copA1*.

Characterization of *copA1* region in pAM330-replicon To analyze the function of the *copA1* region in plasmid replication, six mutations (designated *copA2* to *copA7*) including base-substitution and base-deletion were introduced in the region around *copA1* to construct pVC7N-derived mutants (Table 1; Fig. 4B), and then their transformability into *C. glutamicum* and copy number were examined. We could not obtain transformants of the mutant plasmids containing *copA5* or *copA7*. Mutant plasmids *copA2*, *copA3*, *copA4* and *copA6* could be introduced into *C. glutamicum* and all displayed higher copy numbers than the original pVC7N (Fig. 4A). Among them, the *copA2* mutant plasmid exhibited the highest copy number, 318 ± 56 per chromosome. This plasmid was designated pVC7H2 and contained a mutation from cytosine to guanine at position 1172. These findings indicated that the *copA1* locus is crucial for control of the plasmid copy number from the pAM330-replicon in *C. glutamicum*.

The *copA* region is defined here as the region including bases 1000 to 1300 (Fig. 2B). In this region, no ORF containing more than

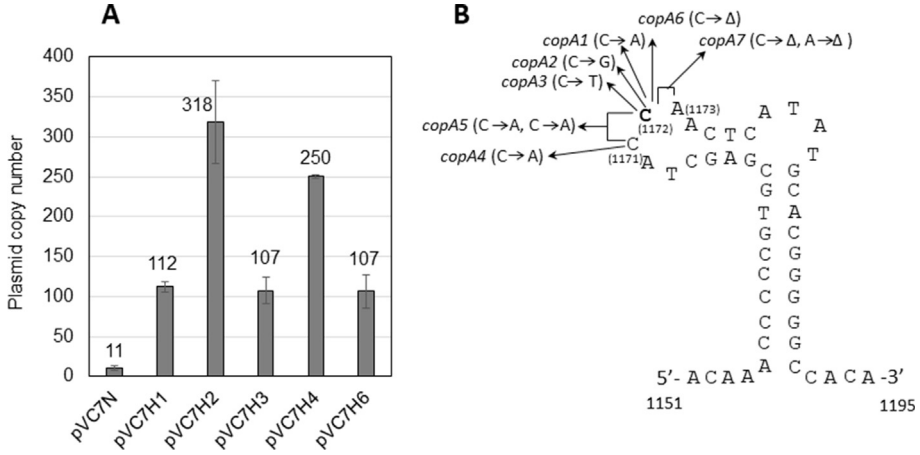

FIG. 4. High copy number (*cop*) mutant plasmids of pVC7N. (A) Plasmid copy number of pVC7N and its mutants in *Corynebacterium glutamicum* 2256L. Copy number is expressed as number of copies of plasmid DNA per copy of chromosomal DNA. Error bars represent ±1 SD (n = 3). (B) Positions of *cop* mutations in a secondary-structure model based on the DNA sequence from position 1151 to 1195 of pVC7N. Each mutation, including base-substitutions (indicated by arrows) or deletions (Δ), was at positions 1171, 1172 and/or 1173. Details of the mutation are given in parentheses.

150 nucleotides was detected. Subsequently, to examine any *trans*-acting factor produced from the *copA* region, four DNA segments consisting of various lengths of this region (Fig. 5A) were cloned into pPK4, which is compatible with pVC7N in *C. glutamicum* 2256L, to construct pPK4-derivatives pPK4-F1R2, pPK4-F1R3, pPK4-F2R1 and pPK4-F3R1 (Table 1). The copy number and/or incompatibility of pVC7N within a single cell harboring each pPK4-derivative were investigated in the presence of Km (but not Cm). All the pPK4 derivatives except pPK4-F1R3 obviously suppressed the copy number of pVC7N (Fig. 5B). In particular, pPK4-F1R2 and pPK4-F3R1 respectively carrying segments F1-R2 and F3-R1 from the *copA* region reduced the copy number of pVC7N severely. From these results, we suggest that the region around the F3-R2 segment, including the *copA1* locus, may produce some *trans*-acting regulator that can suppress the replication of pVC7N. On the other hand, the F1-R3 fragment did not contain enough of this region to regulate the copy number of pVC7N.

Transcript analysis in the *copA* region In previous studies on replication control of pGA1, pCG1 and pCGR2 in *C. glutamicum*, it was reported that a short ctRNA complementary to the leader sequence of *rep*-mRNA is a main regulatory element involved in negative control of *rep* gene expression (17,18). With reference to these studies, we analyzed RNA molecules transcribed *in vivo* from the *copA* region of the pAM330-replicon by RNA-Seq analysis of *C. glutamicum* 2256L harboring pVC7N or pVC7H2. The data obtained for transcriptional organization in the *copA* region revealed that a short RNA is transcribed from the region including the *copA1* locus in the opposite direction from transcription of the putative mRNA for the RepA protein (*repA*-mRNA) (Fig. 2A). This short RNA was estimated to be about 160 nt long, from positions 1200 to 1040 of the DNA sequence in Fig. 2. We performed precise 5′-mapping of the two transcripts (the short RNA and the possible *repA*-mRNA) in the *copA* region by 5′-RACE analysis. Start points of both transcripts were identified as shown in Fig. 2B. The transcriptional start site of the short RNA was mapped as base G at position 1212, only 3 nt upstream of the presumed translational initiation site (TTG triplet) of the *repA* gene on the complementary strand. The transcriptional start

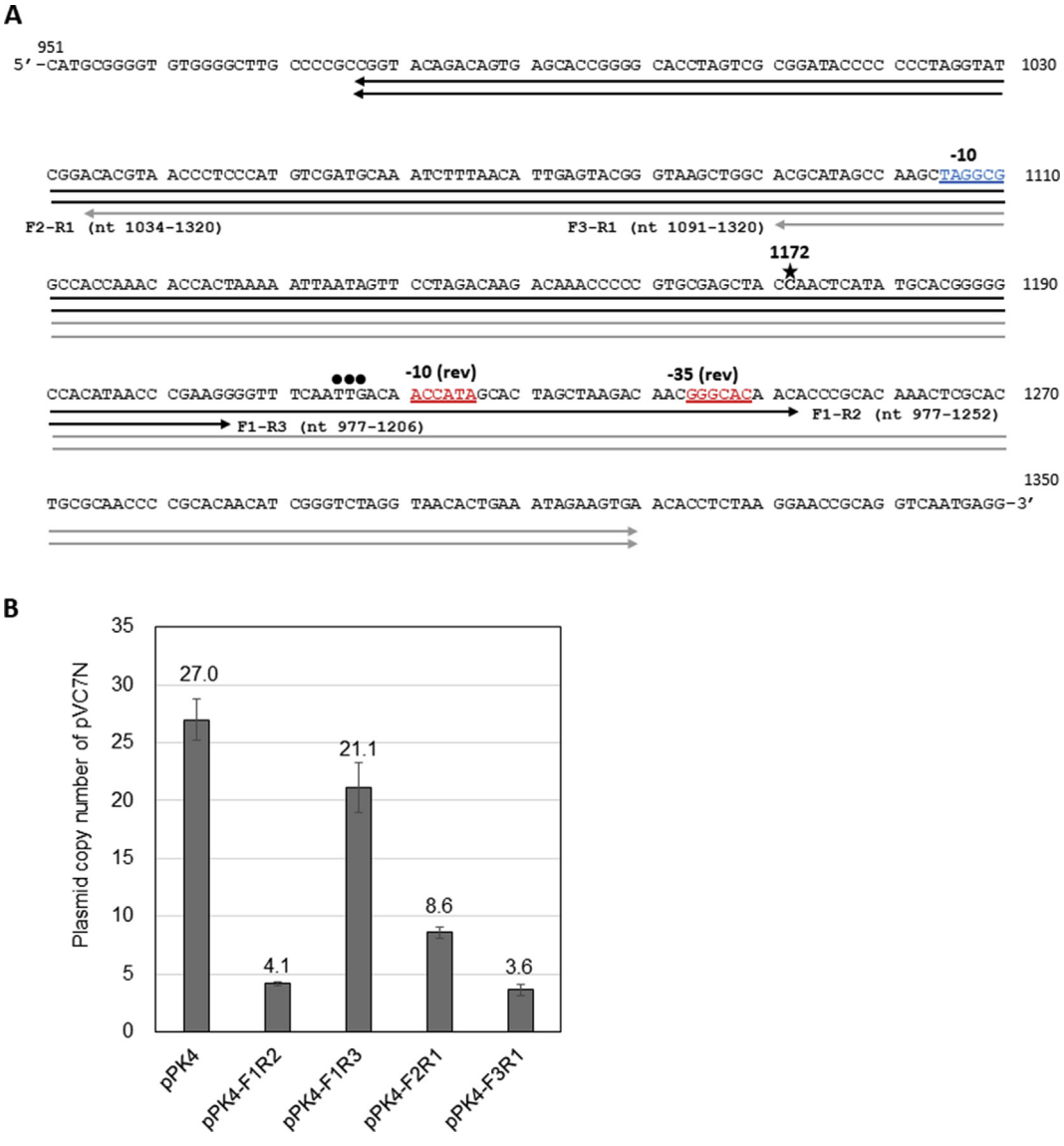

FIG. 5. Effect of segments in the *copA* region on copy number of pVC7N. (A) Each DNA fragment from F1-R2 to F3-R1 (5′- and 3′- end nucleotide positions in each fragment are indicated in parenthesis) was cloned into pPK4 and introduced into *C. glutamicum* 2256L/pVC7N. The presumed initiation codon TTG is indicated with dots and *copA1* mutation at position 1172 is shown with an asterisk. The positions of the putative promoters are shown in red for sRNA1 and blue for *repA* (see text and Fig. 2B for details). (B) Plasmid copy number of pVC7N within each strain harboring the individual DNA fragment shown in panel A in vector pPK4. Error bars represent ±1 SD (n = 3).

point of the *repA*-mRNA was mapped to base C at position 1120. Therefore, the 5′- portion of the short RNA is entirely complementary to the presumed *rep*-mRNA RBS (5′-AAGGGG-3′) (Fig. 2B) which is located 7 nt upstream of the presumed translational initiation codon (TTG) of RepA. Furthermore, in the upstream region of the transcriptional start site of the short RNA, the putative promoter hexamers (5′-GTGCC-3′ at position −35 and 5′-TATGGT-3′ at −10), showing similarity to the *C. glutamicum* σ^A promoter consensus sequences (5′-TTGCC-3′ and 5′-TAN(A/G)(A/C)-3′, respectively (30)) were detected. As a possible promoter sequence for *repA*-mRNA, only 5′-TAGGC-3′ corresponding to the −10 region was detected, 9 nt from the transcriptional start site (we could not find the −35 region of the putative promoter of the *repA* gene) (Fig. 2B).

Subsequently we tried to determine the 3′- terminus of the short RNA by 3′-RACE, but we were not able to do so. However, based on RNA transcripts identified by RNA-Seq analysis, we tentatively concluded that the short RNA (designated sRNA1) is about 170-nt long. The secondary-structure of sRNA1 was predicted using the mfold algorithm (27) (Fig. 6). In the modeled structure, five SL structures (named SL-I to SL-V) could be formed in the RNA from the 5′- to the 3′- end. Mutations *copA1* and *copA2* are on a loop-out region in SL-II. Similar secondary structure of sRNA1 (i.e., five SL-structures) was also predicted by CentroidFold software (28) (Fig. S3). Thus, we suggest that the *copA* region produces sRNA1 of about 170 nt and that this short RNA could regulate plasmid replication.

Characterization of sRNA1 In general, for RNA to act as a regulatory factor, its secondary structure is critical. To evaluate the structure−function relationship of sRNA1, we examined the effect of some other mutations in the *copA* region encoding the putative sRNA1 molecule on the copy number of the cognate pVC7N-derivatives. As Fig. 6 shows, 12 mutants of pVC7N (*mut11* to *mut22*) were constructed and their copy numbers were analyzed in *C. glutamicum* 2256L. Mutations included a single base substitution or deletion by site-directed mutagenesis at arbitrary points in four of the predicted SL structures (SL-II to SL-V in Fig. 6). Mutations *mut11* to *mut14* were introduced in positions affecting SL-II, and *mut15* and *mut16* were in the region between SL-II and SL-III. Mutations *mut17* and *mut18*, *mut19* and *mut20*, and *mut21* and *mut22*, were made in SL-III, SL-IV and SL-V, respectively (Fig. 6). When these plasmid mutants were introduced into *C. glutamicum* 2256L, transformants were obtained for all 12 with comparable efficiencies to transformation with wild-type pVC7N. After isolation of each plasmid from the transformant, the copy number was examined. As a result, significant changes in the copy number were not observed for mutants *mut15* to *mut22*, although pVC7-C1063G (*mut21*) and pVC7-A1131T (*mut17*) displayed a slight increase in the copy number. However, the copy number was considerably increased for all four mutants having individual mutations in the SL-II region (*mut11* to *mut14*) (Fig. S4). This result suggested that SL-II in sRNA1 is particularly important for copy number regulation, as also shown by the high copy number plasmids generated by mutations *copA1*, *copA2*, *copA3*, *copA4* and *copA6* (Fig. 4).

To reconfirm the importance of the possible SL-II structure in copy number control, we examined a compensatory mutation suppressing the *cop* phenotype caused by *mut11* (which was C1183G in the DNA sequence of pVC7N; pVC7-C1183G) in sRNA1. *mut11* changed base G near a bulge in SL-II into C in the sRNA1 sequence, which could change the direction of the top-loop segment of the SL-II structure. In the predicted secondary structure of sRNA1, base G at position 1183 pairs with a C at position 1163 (Fig. 6). Thus, we introduced an additional mutation (G1163C in the DNA sequence of pVC7N) to generate a C1183G/G1163C double mutant (*mut23*). As a result, its copy number was close to that of wild-type pVC7N (Fig. 7). In addition, we examined the effect of deletion of the G at position 1163 on the copy number. The resultant plasmid, pVC7-1163d1183G (*mut24*), had a high copy number, comparable to that of pVC7-C1183G. These results implied that G1163C-substitution could compensate for the conformational change induced by the first (C1183G) mutation in *mut11*, thereby restoring the original function of sRNA1. However, the deletion

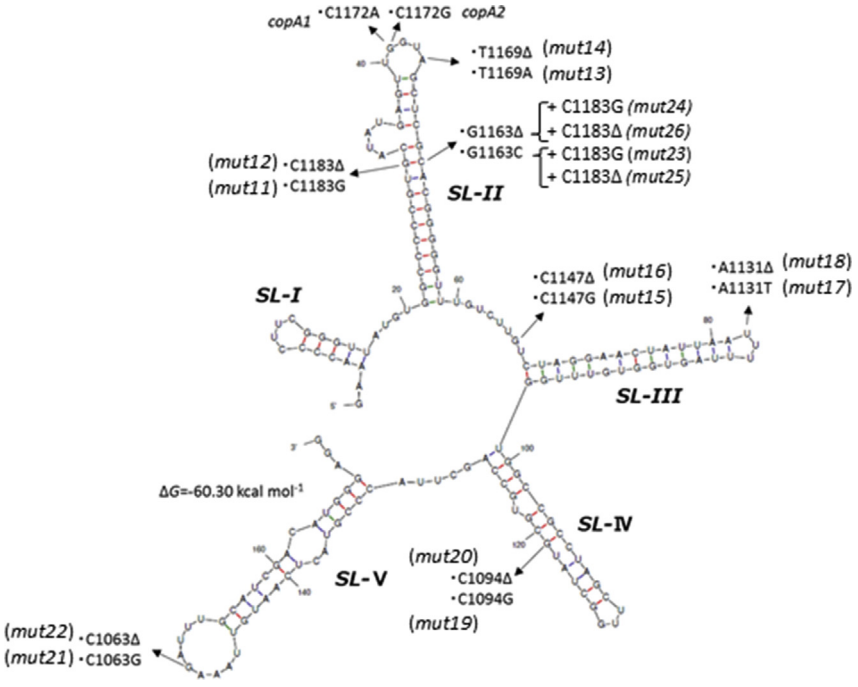

FIG. 6. Predicted secondary structure of sRNA1. Secondary structure of presumed sRNA1 was predicted using mfold software. Five stem-loop structures (SL-I to SL-V) are illustrated from the 5′- to the 3′- end of the RNA. Positions of mutations in this region are indicated based on changes in bases on the DNA sequence, not RNA.

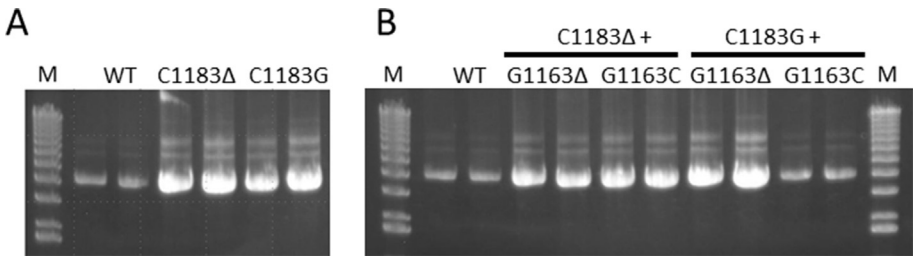

FIG. 7. Effect of mutations at positions of 1163 and 1183 on copy number of pAM330-replicon-containing plasmids. (A) Each plasmid was purified from two independent cultures of *C. glutamicum* 2256L harboring the corresponding plasmid and was electrophoresed on an agarose-gel. WT, C1183G, and C1183Δ indicate pVC7N, pVC7-C1183G (*mut11*) and pVC7-C1183d (*mut12*), respectively. For example, C1183G and C1183Δ means mutation of base C at position 1183 to G and base deletion of C1183, respectively. Lanes marked M show size markers. (B) WT, C1183Δ + G1163Δ, C1183Δ + G1163C, C1183G + G1163Δ and C1183G + G1163C indicate pVC7N, pVC7-1163d1183d (*mut26*), pVC7-1163C1183d (*mut25*), pVC7-1163d1183G (*mut24*) and pVC7-1163C1183G (*mut23*), respectively.

mutant at position 1163 failed to form the functional RNA structure that could restore sRNA1 function impaired in *mut11*. Furthermore, we constructed *mut25* and *mut26*, which contained double mutations C1183Δ/G1163C and C1183Δ/G1163Δ, respectively, and investigated the effect of these mutations on the copy number of each cognate plasmid. These plasmids (pVC7-1163C1183d and pVC7-1163d1183d) both displayed high copy number phenotypes, almost equivalent to that of pVC7-C1183d (Fig. 7B). From these results, we conclude that the SL-II secondary structure is particularly important for the function of sRNA1.

Stability of *cop* mutant plasmids The structural and maintenance stability of the high copy number vectors obtained here was investigated, since such features are also important in useful vectors. The stability of two representative high copy number vectors, pVC7H1 and pVC7H2, was examined in *C. glutamicum* 2256L. In culture without chloramphenicol, pVC7H1 was sustained in almost 100% of cells after about 60 generations, but pVC7H2 was only maintained in about 3% of cells. In the presence of chloramphenicol, both plasmids showed a maintenance stability of 100%. No structural changes were observed in the plasmids in any conditions we tested.

DISCUSSION

A cryptic plasmid pAM330N sequenced here was different in some respects from pAM330 previously reported (5). In addition, it was found that the base T at position 1140 in the sequence of pVC7N is the same as its corresponding base in the sequence of the original pAM330. Thus, although we could not determine whether the major cause of the change was misreading of the DNA sequence in the earlier work, we confirmed that pAM330N and pBL1 are very closely related plasmids (Fig. S2). In previous analysis of ORFs in the autonomous replication region of pAM330 (5), it was indicated that a large ORF comprising 1203 bases encoded an initiator protein. Indeed, the GeneMarkS program for ORF detection (26) predicted an ORF starting with an ATG codon (at positions 1395–1397) and comprising 401 amino acid residues. However, we surmised that in fact RepA is encoded by ORF-1 starting at a TTG codon (at positions 1215–1217) and, therefore, RepA is composed of 461 amino acid residues. The EasyGene program that we employed for ORF prediction and which supported this result uses classical Hidden Markov Models in combination with protein similarity searches (BLASTP sequence analyses) (25). We tried to experimentally determine the N-terminus of RepA through expression of a putative RepA-GFP fusion protein. However, the authentic N-terminus could not be identified because the fusion protein produced was degraded up to the GFP sequence (data not shown). In *C. glutamicum*, translation of some proteins, including transketolase and aconitase, starts from a TTG codon (29,32,33). Since it is generally thought that translation initiated from a TTG codon is inefficient, we speculate that production of RepA is suppressed to a relatively low level. However, intriguingly, the presumed RBS for the *repA* gene is very close to the consensus RBS sequence in *C. glutamicum* (30). From these findings, although the expression of *repA* is generally at a low level, we infer that its expression level could be greatly influenced by the access efficiency of the ribosome to the RBS site on *rep*-mRNA.

In screening for high copy number mutant plasmids derived from pVC7N by spontaneous mutation, only pVC7H1, having mutation *copA1*, was obtained in this trial. However, several mutations that we introduced to the *copA1* locus also caused the elevation of copy number in the cognate plasmids. Through RNA-Seq and genetic analyses, it was shown that the *copA* region encodes a small RNA (sRNA1), and a presumed RBS for RepA was illustrated to be sequestered by the 5′- region of the sRNA1 transcript (Fig. 2B). In the incompatibility test using *copA* DNA segments cloned into pPK4, DNA fragments containing both the sRNA1-coding region and its putative promoter reduced the copy number of pVC7N in cells. However, fragment F1-R3 did not lead to an obvious reduction of the copy number (Fig. 5B). Since F1-R3 does not contain the putative promoter for sRNA1 production no functional sRNA1 is expected to be produced from this fragment, and this is probably why it did not decrease the copy number of pVC7N substantially.

From our results, we suggest that the copy number of pAM330-replicon-containing plasmids can be controlled by the production level of RepA, which is subject to translational efficiency effects through RNA–RNA interaction. Such a control mechanism of copy number by antisense RNA at the translational level was found in RCR plasmids including the Gram-positive bacterial plasmid pC194 and pMV158 families (34,35). In the *C. glutamicum* cryptic plasmid pGA1 (of the pCG1-family), it was suggested that a small ctRNA, which may inhibit translation of the pGA1-*rep* gene, is one of the regulators of plasmid copy number (17). The length of the ctRNA is estimated to be 89 nt and a secondary structure possibly containing two stem-loops is predicted. In addition, antisense-RNA-mediated plasmid copy number control in pCG1 and pCGR2 was reported, and it was indicated that they harbor antisense RNA genes encoding short RNA molecules CrrI (72 nt) and CgrI (73 nt) that are homologous to each other (18). Each of them was also predicted to form two distinctive SL structures. Each small RNA gene in pGA1, pCG1 and pCGR2 has a T-stretch behind the second SL structure, which is a typical feature of a ρ-independent transcriptional terminator. sRNA1 also has a hairpin structure like a typical transcriptional terminator in the second SL structure and it was assumed that this could serve as a terminator (5); however, it may be insufficient for this function, which might explain the somewhat long transcript of sRNA1 (approximately 170 nt).

Mutational analysis of sRNA1 suggested the importance of the SL-II structure for copy number control in cognate plasmids. In

loop-out structures, both 5′-UUCG-3′ in SL-I and 5′-UUGG-3′ in SL-II exhibit the U-turn motif (5′-YUNR-3′, where Y is pyrimidine, U is uracil, N is any nucleoside, and R is purine) (Fig. S5), which is the consensus sequence of a U-turn loop structure acting as a binding-rate enhancer that promotes RNA−RNA pairing (19). In this study, *C. glutamicum* transformants could not be obtained with plasmids including mutation *copA5* or *copA7*. Each of these mutations changes two bases in the important U-turn motif sequence; thus, the effect of this sequence motif in enhancing RNA binding may have been lost in these mutants, which might have caused runaway plasmid replication. The mutated position in *copA1*, *copA2*, *copA3* and *copA6* is base 1172, within the U-turn motif sequence. Interestingly there was a difference in the copy numbers of these four mutant plasmids, with the highest copy number exhibited by *copA2* (mutation C to G at position 1172 in the DNA sequence of pVC7N). Therefore, we suggest that the U-turn motif including base G at position 1172 is of considerable importance for RNA−RNA interactions. In addition, an increase in plasmid copy number caused by *mut11* was suppressed by additional mutation of the nucleotide that is assumed to base-pair with the mutation site of *mut11* in the predicted secondary structure of sRNA1. This result implied the significance of the secondary structure of SL-II in sRNA1, rather than its nucleotide sequence, for its function in copy number control of the pAM330-replicon plasmids.

When we performed homology searches with the DNA sequence (∼100 bases) upstream of the *repA* gene in pAM330N using BLASTN, the corresponding regions from pCG2, pCC1 and pAG3 (all members of the pBL1-family) were detected with >79% identity. In particular, as shown in Fig. 3, high similarity in the sequences of the corresponding regions between pAM330N and pCC1 from *C. callunae* ATCC 15991 (31,36) implies a common mechanism of plasmid copy number control of these two plasmids. However, further biochemical studies will be necessary for a more detailed elucidation of the control mechanism.

Through this work, we obtained high copy number mutant plasmids derived from pAM330N; in particular, pVC7H2 displayed a very high copy number of about 320 per chromosome. No pBL1-family plasmid showing such a high copy number has been reported so far. In previous study, we obtained high copy number plasmids derived from pHM1519 (14), which can coexist with pAM330N, and thus we could further increase target gene dosage by use of both *cop* mutant vectors in *C. glutamicum*. Stability of a plasmid throughout the cultivation process of the host cell is a fundamental requirement in its function as a vector. Plasmid pVC7H1 showed very high stability, but plasmid pVC7H2, which has a higher copy number, was less stable during cultivation. Since the growth rate of *C. glutamicum* harboring pVC7H2 was slower than that of the strain harboring pVC7N, we infer that the shuttle plasmid pVC7H2 may place a metabolic burden on the host cell. *C. glutamicum* has recently attracted increased attention as a host strain for production of useful compounds (37). High copy number plasmids capable of increasing target gene dosage will be beneficial in biotechnological research and for bioindustries employing *C. glutamicum* as a chassis host cell.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jbiosc.2018.10.012.

ACKNOWLEDGMENTS

We are grateful to A. Ando and Y. Iseki for their technical assistance. We thank Drs. T. Nagahiko and Y. Matsuda of Ajinomoto Co. for providing dGFP plasmid and for valuable comments on the experiments, respectively. We thank James Allen, DPhil, from Edanz Group for editing a draft of this manuscript.

References

1. **Yasueda, H.:** Overproduction of L-glutamate in *Corynebacterium glutamicum*, pp. 165−178, in: Anazawa, H. and Shimizu, S. (Eds.), Microbial production - from genome design to cell engineering. Springer, Heidelberg (2014).
2. **Kimura, E.:** L-Glutamate production, pp. 439−464, in: Eggeling, L. and Bott, M. (Eds.), Handbook of *Corynebacterium glutamicum*. Taylor and Francis, Boca Raton (2005).
3. **Liebel, W., Ehrmann, M., Ludwig, W., and Schleifer, K. H.:** Transfer of *Brevibacterium divaricatum* DSM 20297^T, "*Brevibacterium flavum*" DSM 20411, "*Brevibacterium lactofermentum*" DSM 20412 and DSM 1412, and *Corynebacterium lilium* DSM 20137^T to *Corynebacterium glutamicum* and their distinction by rRNA gene restriction patterns, Int. J. Syst. Bacteriol., **41**, 255−260 (1991).
4. **Miwa, K., Matsui, H., Terabe, M., Nakamori, S., Sano, K., and Momose, H.:** Cryptic plasmids in glutamic acid-producing bacteria, Agric. Biol. Chem., **48**, 2901−2903 (1984).
5. **Yamaguchi, R., Terabe, M., Miwa, K., Tsuchiya, M., Takagi, H., Morinaga, Y., Nakamori, S., Sano, K., Momose, H., and Yamazaki, A.:** Determination of the complete nucleotide sequence of *Brevibacterium lactofermentum* plasmid pAM330 and analysis of its genetic information, Agric. Biol. Chem., **50**, 2771−2778 (1986).
6. **Santamaría, R., Gil, J. A., Mesas, J. M., and Martín, J. F.:** Characterization of an endogenous plasmid and development of cloning vectors and a transformation system in *Brevibacterium glutamicum*, J. Gen. Microbiol., **130**, 2237−2246 (1984).
7. **Fernandez-Gonzalez, C., Cadenas, R. F., Noirot-Gros, M. F., Martin, J. F., and Gil, J. A.:** Characterization of a region of plasmid pBL1 of *Brevibacterium lactofermentum* involved in replication via the rolling circle model, J. Bacteriol., **176**, 3154−3161 (1994).
8. **Deb, J. K. and Nath, N.:** Plasmids of corynebacteria, FEMS Microbiol. Lett., **175**, 11−20 (1999).
9. **Tauch, A.:** Native plasmids of amino acid-producing corynebacteria, pp. 57−80, in: Eggeling, L. and Bott, M. (Eds.), Handbook of *Corynebacterium glutamicum*. Taylor and Francis, Boca Raton (2005).
10. **Moriya, M., Matsui, H., Yokozeki, K., Hirano, S., Hayakawa, A., Izui, M., and Sugimoto, M.:** Method of amplifying genes using artificial transposons in coryneform bacteria. United States Patent, US 5804414 (1998).
11. **Takeshita, S., Sato, M., Toba, M., Masahashi, W., and Hashimoto-Gotoh, T.:** High-copy-number and low-copy-number plasmid vectors for LacZ alpha-complementation and chloramphenicol- or kanamycin-resistance selection, Gene, **61**, 63−74 (1987).
12. **Ozaki, A., Katsumata, R., Oka, T., and Furuya, A.:** Functional expression of the genes of *Escherichia coli* in Gram-positive *Corynebacterium glutamicum*, Mol. Gen. Genet., **196**, 175−178 (1984).
13. **Hirano, S., Sugimoto, M., Nakano, E., Izui, M., Hayakawa, A., Yoshihara, Y., and Nakamura, T.:** Methods of producing L-lysine. United States Patent, US 6090597 (1996).
14. **Hashiro, S. and Yasueda, H.:** Plasmid copy number mutation in *repA* gene encoding RepA replication initiator of cryptic plasmid pHM1519 in *Corynebacterium glutamicum*, Biosci. Biotechnol. Biochem., **82**, 2212−2224 (2018).
15. **Sonnen, H., Thierbach, G., Kautz, S., Kalinowski, J., Schneider, J., Pühler, A., and Kutzner, H. J.:** Characterization of pGA1, a new plasmid from *Corynebacterium glutamicum* LP-6, Gene, **107**, 69−74 (1991).
16. **Nešvera, J., Pátek, M., Hochmannová, J., Abrhámová, Z., Bečvářová, V., Jelínková, M., and Vohradský, J.:** Plasmid pGA1 from *Corynebacterium glutamicum* codes for a gene product that positively influences plasmid copy number, J. Bacteriol., **179**, 1525−1532 (1997).
17. **Venkova-Canova, T., Pátek, M., and Nešvera, J.:** Control of *rep* gene expression in plasmid pGA1 from *Corynebacterium glutamicum*, J. Bacteriol., **185**, 2402−2409 (2003).
18. **Okibe, N., Suzuki, N., Inui, M., and Yukawa, H.:** Antisense-RNA-mediated plasmid copy number control in pCG1-family plasmids, pCGR2 and pCG1, in *Corynebacterium glutamicum*, Microbiology, **156**, 3609−3623 (2010).
19. **Franch, T., Petersen, M., Wagner, E. G. H., Jacobsen, J. P., and Gerdes, K.:** Antisense RNA regulation in prokaryotes: rapid RNA-RNA interaction facilitated by a general U-turn loop structure, J. Mol. Biol., **294**, 1115−1125 (1999).
20. **Chinen, A., Kozlov, Y. I., Hara, Y., Izui, H., and Yasueda, H.:** Innovative metabolic pathway design for efficient L-glutamate production by suppressing CO_2 emission, J. Biosci. Bioeng., **103**, 262−269 (2007).
21. **Chinen, A., Yasueda, H., and Nakamura, J.:** Method for producing L-amino acid. United States Patent, US 8012722 B2 (2001).
22. **Hirano, S., Nakamura, J., and Ito, H.:** L-Glutamic acid-producing bacterium and method for production of L-glutamic acid. United States Patent, US 8110381 B2 (2012).
23. **van der Rest, M. E., Lange, C., and Molenaar, D.:** A heat shock following electroporation induces highly efficient transformation of *Corynebacterium glutamicum* with xenogeneic plasmid DNA, Appl. Microbiol. Biotechnol., **52**, 541−545 (1991).

24. **Tsuchida, Y., Kimura, S., Suzuki, N., Inui, M., and Yukawa, H.:** Characterization of a new 2.4-kb plasmid of *Corynebacterium casei* and development of stable corynebacterial cloning vector, Appl. Microbiol. Biotechnol., **81**, 1107–1115 (2009).
25. **Larsen, T. S. and Krogh, A.:** EasyGene - a prokaryotic gene finder that ranks ORFs by statistical significance, BMC Bioinformatics, **4**, 21 (2003).
26. **Besemer, J., Lomsadze, A., and Borodovsky, M.:** GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions, Nucleic Acids Res., **29**, 2607–2618 (2001).
27. **Zuker, M.:** Mfold web server for nucleic acid folding and hybridization prediction, Nucleic Acids Res., **31**, 3406–3415 (2003).
28. **Sato, K., Hamada, M., Asai, K., and Mituyama, T.:** CENTROIDFOLD: a web server for RNA secondary structure prediction, Nucleic Acids Res., **37**, W277–W280 (2009).
29. **Kalinowski, J.:** The genomes of amino acid-producing Corynebacteria, pp. 37–56, in: Eggeling, L. and Bott, M. (Eds.), Handbook of *Corynebacterium glutamicum*. Taylor and Francis, Boca Raton (2005).
30. **Pfeifer-Sancar, K., Mentz, A., Rückert, C., and Kalinowski, J.:** Comprehensive analysis of the *Corynebacterium glutamicum* transcriptome using an improved RNAseq technique, BMC Genomics, **14**, 888 (2013).
31. **Venkova-Canova, T., Pátek, M., and Nešvera, J.:** Characterization of the cryptic plasmid pCC1 from *Corynebacterium callunae* and its use for vector construction, Plasmid, **51**, 54–60 (2004).
32. **Ikeda, M., Kamada, N., Takano, Y., and Nakano, T.:** Molecular analysis of the *Corynebacterium glutamicum* transketolase gene, Biosci. Biotechnol. Biochem., **63**, 1806–1810 (1999).
33. **Krug, A., Wendisch, V. F., and Bott, M.:** Identification of AcnR, a TetR-type repressor of the aconitase gene *acn* in *Corynebacterium glutamicum*, J. Biol. Chem., **280**, 585–595 (2005).
34. **Alonso, J. C. and Tailor, R. H.:** Initiation of plasmid pC194 replication and its control in *Bacillus subtilis*, Mol. Gen. Genet., **210**, 476–484 (1987).
35. **López-Aguilar, C., Ruiz-Masó, J. A., Rubio-Lepe, T. S., Sanz, M., and del Solar, G.:** Translation initiation of the replication initiator *repB* gene of promiscuous plasmid pMV158 is led by an extended non-SD sequence, Plasmid, **70**, 69–77 (2013).
36. **Sandoval, H., Aguilar, A., Paniagua, C., and Martín, J. F.:** Isolation and physical characterization of plasmid pCC1 from *Corynebacterium callunae* and construction of hybrid derivatives, Appl. Microbiol. Biotechnol., **19**, 409–413 (1984).
37. **Lee, J.-Y., Na, Y.-A., Kim, E., Lee, H.-S., and Kim, P.:** The Actinobacterium *Corynebacterium glutamicum*, an industrial workhorse, J. Microbiol. Biotechnol., **26**, 807–822 (2016).