

# Sequence Variations in the Non-Coding Sequence of CTX Phages in *Vibrio cholerae*

Eun Jin Kim<sup>1,2</sup>, Hyun Jin Yu<sup>1,2</sup>, and Dong Wook Kim<sup>1,2\*</sup>

<sup>1</sup>Department of Pharmacy, College of Pharmacy, Hanyang University, Ansan 15588, Republic of Korea

<sup>2</sup>Institute of Pharmacological Research, Hanyang University, Ansan 15588, Republic of Korea

Received: April 7, 2016  
Revised: May 2, 2016  
Accepted: May 9, 2016

First published online  
May 9, 2016

\*Corresponding author  
Phone: +82-31-400-5806;  
Fax: +82-31-400-5958;  
E-mail: dongwook@hanyang.ac.kr

pISSN 1017-7825, eISSN 1738-8872

Copyright© 2016 by  
The Korean Society for Microbiology  
and Biotechnology

This study focused on the variations in the non-coding sequences between *ctxB* and *rstR* of various CTX phages. The non-coding sequences of CTX-1 and CTX-cla are phage type-specific. The length of the non-coding region of CTX-1 and CTX-cla is 601 and 730 nucleotides, respectively. The non-coding sequence of CTX phage could be divided into three regions. There is a phage type-specific Variable region between two homologous Common regions (Common regions 1 and 2). The non-coding sequence of RS1 element is similar to CTX-1 except that Common region 1 is replaced by a short RS1-specific sequence. The non-coding sequences of CTX-2 and CTX-cla are homologous, indicating the non-coding sequence of CTX-2 is derived from CTX-cla. The non-coding region of CTX-O139 is similar to CTX-cla and CTX-2; however, it contains an extra phage type-specific sequence between Common region 2 and *rstR*. The variations in the non-coding sequences of CTX phages might be associated with the difference in the replication efficiency and the directionality in the integration into the *V. cholerae* chromosome.

**Keywords:** *V. cholerae*, CTX phage, non-coding sequence

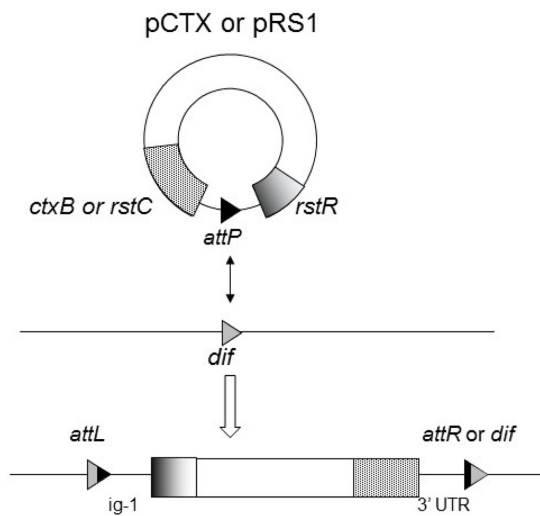
## Introduction

The CTX phage (cholera toxin phage) that provides the cholera toxin gene to *Vibrio cholerae* strains, making them toxigenic bacteria, has been classified into two major types, classical type CTX phage (CTX-cla) and El Tor type CTX phage (CTX-1) [10]. The CTX phage consists of 10 genes, *rstR*, *rstA*, *rstB*, *psh*, *cep*, *orfU*, *ace*, *zot*, *ctxA*, and *ctxB*. Whereas the genes from *rstA* to *ctxB* are homologous between two phages, *rstR* is phage type-specific and has been used to distinguish two phages [11]. *ctxB* is homologous in two phages; however, two SNPs (single nucleotide polymorphisms, 115<sup>th</sup> and 203<sup>rd</sup> nucleotides) have been noticed as phage type-specific [20]. Although the DNA sequences of the CTX phage genes have been used to discriminate different CTX phages, the non-coding sequence between *ctxB* and *rstR* has not been extensively studied.

Atypical El Tor *V. cholerae* strains that harbor the mosaic CTX phages between the classical and El Tor CTX phages have emerged since 1991 and, currently, most of the clinical

isolates of *V. cholerae* are atypical El Tor variants that contain the mosaic CTX phage [19]. Various mosaic CTX phages have been classified among the strains belonging to Wave 2 and 3 within the 7<sup>th</sup> cholera pandemic [17]. CTX-2 in Wave 2 strains contained classical *rstR* and *ctxB*, while the rest of the phage genome is similar to CTX-1. CTX-3 ~ CTX6 in Wave 3 strains contained a CTX-1-like sequence except that they contain classical *ctxB* [12].

CTX phage is a single strand DNA phage that undergoes a replicative form double strand circular DNA, pCTX, during replication [15]. The integration of CTX phage genome on the *V. cholerae* chromosome is mediated by the host Xer recombination machinery [3]. XerC and XerD recombinases originally mediate the chromosomal dimer resolution after the chromosomal DNA replication via a 28 bp site, *dif* (deletion-induced filamentation), and CTX phage exploits XerCD for integration in the chromosomal DNA [3]. A molecular mechanism that explains the integration of pCTX on the *dif* site has been demonstrated [2]. The *attP* sequence (or XerCD binding sequence) is



**Fig. 1.** Integration of pCTX or pRS1 on the chromosome of *V. cholerae*.

Integration of pCTX and pRS1 is mediated by *attP* (black triangle) on pCTX (or pRS1) and the *dif* site on the chromosome (grey triangle). The non-coding sequence between *rstR* and *ctxB* (or *rstC* in the case of pRS1) is split into *ig-1* and 3' UTR after the integration. Next, pCTX or pRS1 is integrated via a similar process on the *attR* (or new *dif* site) next to CTX or RS1.

located in the non-coding sequence of pCTX that spans from *ctxB* and *rstR* of the circular pCTX phage genome [16].

The *attL* and *attR* sites, generated by the integration of the CTX phage, have been described as the “end repeat” empirically [7]. As shown in Fig. 1, the integration of pCTX in the chromosome via *attP* and the *dif* sequence splits the non-coding sequence of pCTX into two parts; (i) *ig-1* (intergenic sequence-1, previously used to describe the sequence between *attL* and *rstR*), and (ii) 3' UTR (untranslated region) sequence between *ctxB* and *attR* (Fig. 1). This process is similar in the integration of pRS1.

*V. cholerae* strains have two chromosomes (chromosome 1 of about 3 M bp and chromosome 2 of about 1 M bp) and the CTX phages can be integrated in either of the two

chromosomes or both [8]. It was suggested that the CTX-2 ~ CTX-6 could have been generated in a strain that contained different CTX phages on each chromosome; for example, a strain contained the CTX-1 in one chromosome and the CTX-cla in the other. It was demonstrated that CTX-3 ~ CTX-6 could be generated in a strain, V212-1, which contained CTX-1 on chromosome 1 and CTX-2 on chromosome 2 [12].

From the analysis of the sequences of mosaic CTX-3 ~ CTX-6 phages produced in V212-1 by the inter-chromosomal recombination, the generation of pCTX phages with variant non-coding sequences was noticed [12]. These variant sequences contained a non-coding sequence of CTX-2 on chromosome 2, which could be distinguished from the non-coding region of CTX-1 on chromosome 1. Sequence analyses of the non-coding region of CTX prophages and pCTXs produced in various *V. cholerae* strains showed that CTX-cla and CTX-1 had a major different region in the middle of the non-coding sequence that was flanked by Common regions. The phage type-specific regions in the non-coding sequences of pRS1 and pCTX-O139 have also been identified in this study.

## Materials and Methods

### Non-Coding Sequences of pCTXs and pRS1

The non-coding sequences of CTX-1 prophage in two Wave 1 El Tor strains, N16961 and V212-1, were obtained from the genome sequence data base under the accession number of AE003852/AE003853 and ERS013132, respectively [17]. *V. cholerae* serogroup O1 biotype El Tor strain N16961 contains a CTX-1:RS1 array on chromosome 1. By comparison of the CTX phage sequence of N16961 with a strain that does not contain any element on chromosome 1, such as PM4 (GenBank Accession No. KF471410), and a pCTX-1 (GenBank Accession No. KF664569) produced from strain V212-1, the non-coding sequence of pCTX-1 could be deduced (Table 1, GenBank Accession No. KU877492, Fig. 2) [12]. Similarly, the non-coding sequence of the replicative form of the satellite phage RS1, pRS1, could also be deduced (GenBank

**Table 1.** Non-coding sequences of the replicative form of CTX phages and RS1.

Non-coding sequences	GenBank Accession No.
pRS1 non-coding sequence	KU877491
pCTX-1 non-coding sequence	KU877492
pCTX-1-hybrid non-coding sequence	KU877493
pCTX-2 non-coding sequence	KU877494
pCTX-cla non-coding sequence of chromosome 2 of strain O395	KU877495
pCTX-O139 non-coding sequence	KU877496
pCTX-3 non-coding sequence	KU877497

	XerC	XerD
<i>dif1</i>	TGTTCCCT AGTGCGCATTATGTATGTTATGTTAAATTA	
<i>attL</i> <sup>CTX-1</sup>	TGTTCCCT AGTGCGCATTATGTGGCGCGGCATTATGTT	
<i>attR</i> <sup>CTX-1</sup>	ACATCCCT <b>T</b> AGTGCG <b>T</b> ATTATGTATGTTATGTTAAATTA	
<i>attP</i> <sup>CTX-1</sup>	ACATCCCT <b>T</b> AGTGCG <b>T</b> ATTATGTGGCGCGGCATTATGTT	
	XerC	XerD
<i>dif2</i>	AACAC TTAATGCGCATTACGTGCGTTATGTTAAATTA	
<i>attL</i> <sup>CTX-cla</sup>	AACAC TTAATGCGCATTAT <b>T</b> GTGGCGCGGCATTATGTT	
<i>attR</i> <sup>CTX-cla</sup>	ACATCC <b>T</b> TAGTGCG <b>T</b> ATTACGTGCGTTATGTTAAATTA	
<i>attP</i> <sup>CTX-cla</sup>	ACATCC <b>T</b> TAGTGCG <b>T</b> ATTATGTGGCGCGGCATTATGTT	
<i>attL</i> <sup>CTX-2</sup>	AACAC TTAATGCGCATTATGTGGCGCG <b>T</b> CATTATGTT	
<i>attR</i> <sup>CTX-2</sup>	ACATCC <b>T</b> TAGTGCG <b>T</b> ATTACGTGCGTTATGTTAAATTA	
<i>attP</i> <sup>CTX-2</sup>	ACATCC <b>T</b> TAGTGCG <b>T</b> ATTATGTGGCGCG <b>T</b> CATTATGTT	

**Fig. 2.** Extraction of *attP* sequences of pCTX-1, pCTX-cla, and pCTX-2 from the integrated CTX phage sequences.

The *attP*<sup>CTX</sup> sequence could be deduced from the *attL* and *attR* sequences by comparing with *dif* sequences of each chromosome. SNPs unique to each pCTX are shown in bold. The non-coding sequence between *rstR* and *ctxB* (or *rstC*) of pCTX or pRS1 could be deduced from the integrated CTX and RS1.

Accession No. KU877491).

Wave 2 strain B33 contains a tandem repeat of CTX-2 on chromosome 2 (GenBank Accession No. GQ485644) [13]. By comparing the *dif* sequence of chromosome 2 of a strain that does not contain any element on chromosome 2 (GenBank Accession No. FJ449754, strain IB4122) with the CTX-2 sequence of B33, the non-coding sequence of the replicative form of CTX-2, pCTX-2, could be deduced (GenBank Accession No. KU877494) [13]. The DNA sequences of the non-coding region of pCTX-1 and pCTX-2 are shown in Fig. 3.

A classical biotype strain, O395, contains a truncated classical CTX followed by a full-length classical CTX phage on chromosome 1 and a classical CTX phage on chromosome 2 [13]. The non-coding sequences of the CTX-cla prophage in two classical biotype strains, O395 and GP8, were obtained from the genome sequence information under the accession number CP000626/CP000627 and ERS013128, respectively [17]. Since the non-coding sequence of pCTX-cla from the CTX-cla's on chromosome 1 could be a mix of two variant classical phages, the non-coding sequences of CTX-cla on chromosome 2 could be considered the authentic non-coding sequence of pCTX-cla (GenBank Accession No. KU877495).

#### pCTX-O139 Intergenic Sequence

The intergenic sequence of pCTX-O139 was deduced from the

intergenic sequence of integrated CTX-O139 in two O139 strains, AR196157 and AR9954 [14]. The intergenic sequence was determined by PCR and sequencing with the primer set *ctxBF* (AGA TAT TTT CGT ATA CAG AAT CTC TAG) and *rstRcalR* (TGG CAA CAA AGC ACA TTA AAG A) [18]. A previously reported intergenic sequence (GenBank Accession No. AF110029) of an O139 strain, AS207, was also analyzed [4]. The full DNA sequence of the non-coding sequence of CTX-O139 is shown in Fig. 3.

#### pCTX-1-Hybrid Intergenic Sequence

The mosaic pCTX phage generated by the inter-strand recombination between CTX-1 on chromosome 1 and CTX-2 on chromosome 2 in strain V212-1 usually contained the non-coding sequence of CTX-1 on the chromosome. However, a variant of the mosaic CTX phage that contained the mixed non-coding sequences of CTX-1 and CTX-2 was identified and designated as pCTX-1-hybrid.

## Results

### Comparison of the Non-Coding Sequences of CTX-1, CTX-2, and CTX-cla

The non-coding sequences of prophages CTX-1 in strains N16961 and V212-1, CTX-2 in B33, and CTX-cla in O395 and GP8 were deduced from the genome sequence data, as shown in Fig. 2. The non-coding sequences of CTX-1 in N16961 and V212-1 were identical and the CTX-cla's in strains O395 and GP8 contained the same non-coding sequences. These non-coding sequences were compared with the non-coding sequence of the replicative form pCTX-1 (GenBank Accession No. KF664579) [12].

The non-coding sequence of CTX-1 prophage in strain N16961 was identical to that of pCTX-1, which implies that the DNA sequence of the non-coding sequence is not altered by the integration or excision process [1]. The length of the non-coding sequence of CTX-1 was 601 bp (between the termination codons of *rstR* and *ctxB*).

Two CTX-2 prophages in the tandem repeat array on chromosome 2 of strain B33 had identical sequence not only in the ORFs of the CTX phage genome but also in the non-coding sequence. The non-coding sequences of CTX-2 and CTX-cla were homologous and the length was 730 bp (Fig. 3). There are four SNPs between non-coding sequences of CTX-2 and CTX-cla at positions 131, 213, 330, and 432, and the 131<sup>st</sup> position belongs to the XerD binding site (Fig. 2). The non-coding sequences of the replicative form of pCTX-2 and pCTX-cla are not available since they have not been experimentally demonstrated yet; however, they could be assumed to be identical to the non-coding sequence of CTX-2 and CTX-cla prophages, respectively, as

## Common region 1

```

1
pCTX-1 GATATAAAAA AGCCACCTC AGTGGGCTTT TTTGTGGTTC GATGATGAGA AGCAACCGTT TTGCCCAAAC ATGTATTACT GCAAGTATGA TGTTTTTATT 100
pCTX-2 GATATAAAAA AGCCACCTC AGTGGGCTTT TTTGTGGTTC GATGATGAGA AGCAACCGTT TTGCCCAAAC ATGTATTACT GCAAGTATGA TGTTTTTATT

101
pCTX-1 CCACATCCTT AGTGCCTATT ATGTGGCGCG GCATTATGTT GAGGGGCAGT CGTCAGTACC ATTGCGCCAG CACTGACGGC CTCACTTGCA GCAGAACGTG 200
pCTX-2 CCACATCCTT AGTGCCTATT ATGTGGCGCG TCATTATGTT GAGGGGCAGT CGTCAGTACC ATTGCGCCAG CACTGACGGC CTCACTTGCA GCAGAACGTG

201
pCTX-1 GGCAGCTGC TGAATCGTTC TGCAAGAGTG AGCCCGTAAC ATAATGGCGT ATAATACGCA TTAAGGCGGT 270
pCTX-2 GGCAGCTGC TGAATCGTTC TGCCAAAGTG AGCCCGTAAC ATAATGGCGA GTAATACGCA TTAAGGCGGT

```

## Variable regions

```

pCTX-1 ATGTCATTTT GGTATGTCAA AAATGACATA ATTCGATTTA TTCTGATTC AGCCGTCGCG CGCAGTCATC AGCTTCGCTG ATGCGAGGAG ACGGAATTTT
TACAGGTTCT ATTGAGACAG CGGCCGCTGT GAGCTTAATT GTCTCA (136 nt)

pCTX-2 AACTCAGCCC CGCAGGGACT AGACCTAACG TTAGGCTCAG CGCTCGCGC TCTGATGCTA CTGCATATCC AAAGCTGCTT TAGCACTCGC AGAAGTTCCG
TTGATTGCTC AAGCGTTCCC GTCAGTGAAA TGATCCTCTT TCTGATAGCG CCAGAAAAAA CTCCTTCGT CCTGCCAAGC CCATTGGAA GTCTCAGCAC
ACGCAGAGGG TAACAGCATT TGTCATGGAT ACGTTCAGCG CCCAAGGCGC GCGAGAGTC GAGCAAG (267 nt)

```

## Common region 2

```

1
pCTX-1 CCTCT.ATAC TGCACAGCG GCAGGTGAGA ACATAAGCGA CGTAGCGTGC GGAGTCGCGT TGTTAGAGCC TGTCGCTGT GGTAGACCCC CGTCTAGTAT 100
pCTX-2 CCTCTTATAC TGCACAGCG GCAGGTGAGA ACATAAGCGA CGTAGCGTGC GGAGTCGCGT TGTTAGAGCC TGTCGCTGT GGTAGACCCC CGTCTAGTAT

101
pCTX-1 TACGGGGGTA AATCCACAG AGCCTGTGAC ACTCACCTTG TATTGCAAG CGTAGCGCGC CAGTGTTTGA GCGCTAGCGA GTCTTG... ... 193
pCTX-2 TACGGGGGTA AATCCACAG AGCCTGTGAT ACTCACCTTG TATTGCAAG CGTAGCGCGC CAGTGTTTGA GCGCTAGCGA GTCTTTGAT TAC

```

## O139 unique sequence

```

1
TTCATACCTC AAGCGTTCAC TTTTGGACCT CGAAGCGGTT CGGGTTCCTG TTACACCGA ACTAAGTCCC GAATTCAGGA CTTTACTTT GTTAGTTCTT 100
CTGTTTTGTC GCTGAAAAAC TTTGAAGCA TGTTGTTTT TTGCTTTCAA TAGGGCTTTA CGGAATCTCT CGCCATTTC ATTTTGTTAA AATG (194 nt)

```

**Fig. 3.** The non-coding sequences of pCTX-1, pCTX-2, and pCTX-O139.

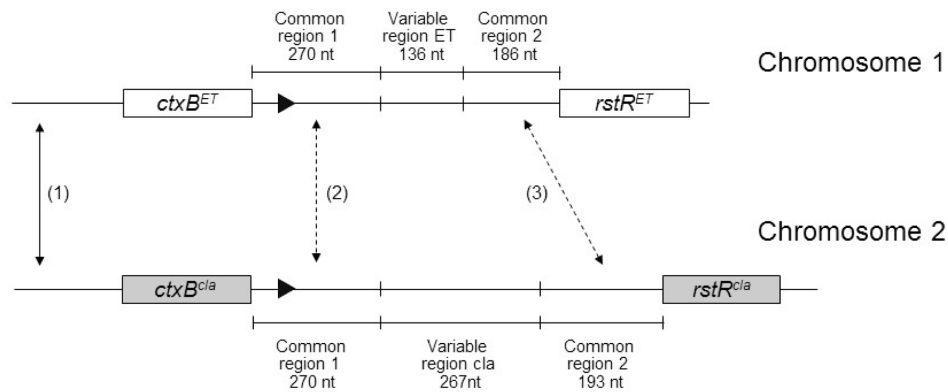
Common region 1 (the first 270 nucleotides after the termination codon of *ctxB*) and Common region 2 (186 nucleotides of pCTX-1 and 193 nucleotides of pCTX-2 before the termination codon of *rstR*) are aligned together, whereas the Variable region (136 nucleotides of pCTX-1 and 267 nucleotides of pCTX-2, shaded) between the common regions are shown individually (shaded). The non-coding sequence of pCTX-O139 contains the same sequences as pCTX-2 and additional 194 nucleotides instead of 11 nucleotides at the end of Common region 2 of pCTX-2 (Italic and underlined). The different nucleotides between pCTX-1 and pCTX-2 are shown in bold (positions 131, 224, 226, 250, and 251 in Common region 1 and positions 6 and 130 in Common region 2). Two nucleotides (positions 131 and 213 in Common region 1) underlined in Common region 1 are different nucleotides between pCTX-2 and the pCTX-cla (both positions are G in pCTX-cla).

shown in CTX-1.

Alignment of the non-coding sequences of CTX-1 and CTX-cla (CTX-2) showed that the non-coding sequence of CTX could be divided into three parts. The first 270 bp (from the termination codon of *ctxB*) are homologous between CTX-1 and CTX-cla, while five SNPs were found (Fig. 3). This part can be designated as Common region 1 of the non-coding sequence. Next to Common region 1, a Variable region with a different length is located. The Variable regions are 136 bp and 267 bp long in CTX-1 and CTX-2, respectively. There is no sequence similarity between the two Variable regions, and therefore, they could not be aligned (Fig. 3). The Variable region is followed by another

Common region (Common region 2) of 185 and 193 bp long in CTX-1 and CTX-2, respectively. Next to Common region 2 is the termination codon of *rstR* (*rstR*<sup>El Tor</sup> in CTX-1 and *rstR*<sup>cla</sup> in CTX-cla). The difference in length of Common region 2 is due to a T addition between the 5<sup>th</sup> and 6<sup>th</sup> nucleotides in CTX-2, and the last nucleotide G of CTX-1 is replaced with TTGATTAC in CTX-2 (Fig. 3).

The non-coding sequence of RS1 element was 511 bp long and contained a similar sequence as CTX-1. Although the Variable region and Common region 2 of the non-coding sequence of RS1 are identical to those of CTX-1, the first 90 bp of Common region 1 are missing in the non-coding sequence of RS1, and nine SNPs in the next 15 bp (Table 1).



**Fig. 4.** Schematic diagram of generation of pCTX-1-hybrid.

The first homologous recombination between CTX-1 on chromosome 1 and CTX-2 on chromosome 2 occurs in any part of CTX phages (recombination position (1)) and the second recombination can occur between Common region 1 or Common region 2. When the recombination occurs on Common region 1 (recombination position (2)), the CTX phage produced on chromosome 1 contains the same intergenic sequence as CTX-1. However, the CTX phage produced on chromosome 1 contains an intergenic sequence similar to CTX-2, when the recombination occurs on Common region 2 (CTX-1-hybrid, recombination position (3)).

#### Non-Coding Sequence of pCTX-O139

The non-coding sequence of CTX-O139 was deduced from the PCR product of the primer set *ctxB*/*rstRcalR* from two O139 serogroup strains [13]. This sequence includes part of the non-coding sequence (between *ctxB* and *attR*) of the preceding CTX prophage, not the CTX-O139 origin sequence. However, this sequence is identical among CTX-1, CTX-cla, and CTX-2, and it was assumed that CTX-O139 also contained the same sequence. The non-coding sequence of pCTX-O139 was similar to the non-coding sequence of pCTX-2 and pCTX-cla, since it contained Common regions 1 and 2 and the Variable region of pCTX-2 (Fig. 3). However, the non-coding sequence of pCTX-O139 contained additional 194 nucleotides instead of the last 11 nucleotides of Common region 2 of pCTX-2. These additional nucleotides in the non-coding sequence of pCTX-O139 was also reported previously [4].

#### Non-Coding Sequence of pCTX-1-Hybrid

CTX-3 ~ CTX-6 that are found in Wave 3 strains have been shown to be generated by an inter-strand homologous recombination between two different CTX prophages as shown in an example of strain V212-1 [12]. The recombination could occur between CTX-1 (followed by an RS1) on chromosome 1 and the first CTX-2 of a tandem repeat array on chromosome 2 in V212-1. The recombination occurs anywhere upstream of *ctxB*s of two phages (position (1) in Fig. 4) and downstream of *ctxB*s to generate a new mosaic CTX phage genome that contains classical *ctxB* on chromosome 1. When the second recombination occurs

downstream of *ctxB*, the recombination can occur via Common region 1 or 2 (recombination position (2) or (3) in Fig. 4, respectively). When the recombination is mediated via Common region 1, the newly generated CTX phage should have the same non-coding sequence as CTX-1. (There are five SNPs between the Common regions 1 of CTX-1 and CTX-2; therefore, five minor variants could be generated via these five SNPs, Fig. 3.) For example, the pCTX-3 found in Wave 3 strain 01.07.VP (IB4122) contains the same non-coding sequence as CTX-1 in the Wave 1 strain N16961 (Table 1), which implies that this pCTX-3 was generated by recombination between the Common regions 1 of CTX-1 and CTX-2.

However, a hybrid non-coding sequence could be generated when the recombination occurred via Common region 2, and the CTX phage generated by this recombination should contain Common region 1, the Variable region of pCTX-2, and a hybrid of the Common regions 2 of CTX-1 and CTX-2. In this hybrid non-coding sequence, the length is 723 bp (7 bp shorter than CTX-2 since the 6<sup>th</sup> nucleotide "T" is from CTX-2 and the TTGATTAC sequence at the end of CTX-2 is replaced with G of CTX-1) (Fig. 3).

The recombination via Common region 1 is more likely to happen, since the recombination occurs in the homologous region of CTX phage (*ctxB* and Common region 1). The recombination position (3) in Fig. 4 is less likely because Common region 2 is only 193 bp long and flanked by two diverse sequences (Variable region in the non-coding sequence and *rstR*). However, this recombination does occur via Common region 1, and pCTX-1-hybrid was indeed

generated. The pCTX-1-hybrid could be integrated on the genome of the transduced recipient strain [12].

### Discussion

The molecular mechanism of the integration of pCTX phage genome via the *dif* site in *V. cholerae* chromosomes has been demonstrated [2, 16]. The excision of CTX phage has not been documented extensively, since the replication mechanism of CTX phage is unique among the lysogenic phages [6]; however, the excision of intact CTX phage genome has been recently demonstrated, which can leave the authentic *dif* site on the *V. cholerae* chromosome [9, 12]. We deduced non-coding sequences of pCTX-cla, pCTX-1, pCTX-O139, and pRS1.

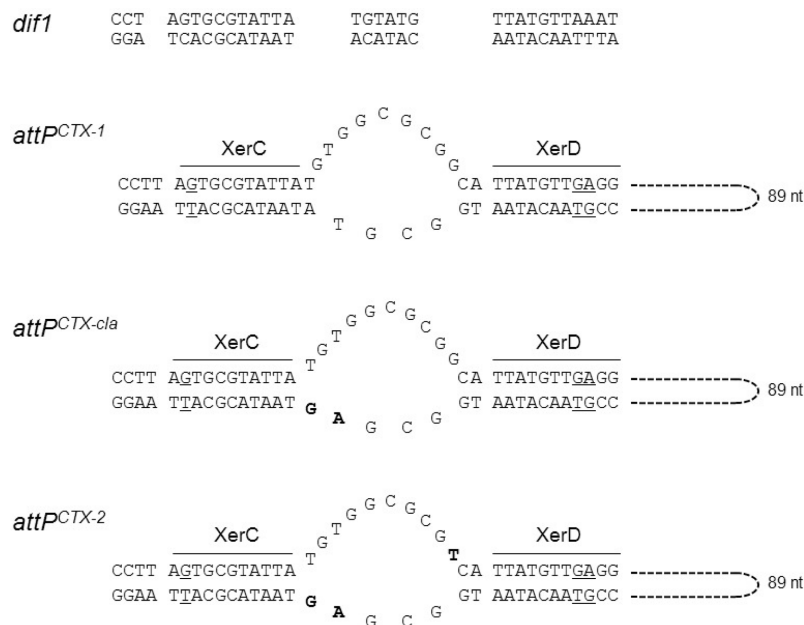
The single-stranded CTX phage forms a stem-loop structure for integration in the *V. cholerae* chromosomes, and the stem-loop structure sequence is located entirely on Common region 1 of the non-coding sequence [2]. The XerC binding site is located at the positions 111–121 and 252–262 of Common region 1, while the XerD binding site is at 134–144 and 234–244. During the integration of pCTX, the non-coding sequence splits into two parts. The first half of Common region 1 (from the first nucleotide to the

XerCD binding site) becomes the 3' UTR and the rest of the non-coding sequence constitutes *ig-1*. There are three SNPs that are located in the bulge of the stem-loop structure of the non-coding sequence of CTX-1, CTX-2, and CTX-cla and a proposed stem-loop structure is shown in Fig. 5 [16].

The non-coding sequence of pCTX-cla has not been described because the replicative form of pCTX-cla has not been experimentally revealed [5]. However, the non-coding sequence of pCTX-cla could be deduced from the integrated CTX-cla by combining the *ig-1* sequence and 3' UTR sequence in a strain that harbors CTX-cla.

While we have shown the generation of pCTX-3 ~ pCTX-6 from V212-1 that contained CTX-1 on chromosome 1 and a tandem repeat of CTX-2 on chromosome 2, we have noticed that a derivative of each pCTX was also generated [12]. The derivative of each mosaic CTX contained the same CTX genome of the newly generated mosaic CTX phage, but contained a different non-coding sequence. We found that the non-coding sequences of pCTX-1 on chromosome 1 and pCTX-2 on chromosome 2 are different.

Previously, the non-coding sequence between the *attR* site and the termination of *rstR* has been named as *ig-1* [4]. The molecular mechanism of the integration of pCTX on the chromosomes of *V. cholerae* via the XerCD recombination



**Fig. 5.** Diagram of the XerC and XerD binding site structure of the single-stranded pCTX's.

*attP* sequences of pCTX-1, pCTX-cla, and pCTX-2 aligned with the *dif* sequence of chromosome 1. The single-stranded CTX phage uses a stem-loop structure to form functional XerC/XerD binding sites for integration in the *V. cholerae* chromosome [16]. *attP<sup>CTX-cla</sup>* and *attP<sup>CTX-2</sup>* contain 2 and 3 SNPs compared with *attP<sup>CTX-1</sup>*. The variant nucleotides (position 131, 250, and 251 of Common region 1) among pCTX-1, pCTX-cla, and pCTX-2 are shown in bold.

system on the *dif* sequence has been unveiled and we now understand the non-coding sequence between *ctxB* and *rstR* on pCTX splits into the *ig-1* and the 3' UTR sequences next to *ctxB* of the prophage.

The non-coding sequences of CTX-2 and CTX-cla are homologous, containing the same Variable region. Although most of the CTX genome of CTX-2 is identical to CTX-1, the *ctxB*, non-coding sequence, and *rstR* of CTX-2 are of CTX-2, which indicates that CTX-2 was perhaps generated by recombination between CTX-1 and CTX-cla.

There have been no experimental evidence of what the functions of the various non-coding sequences in different types of CTX phages are. The CTX phage replication initiation/termination point is shown to be located in Common region 2 [4]; however, the roles of Common region 1 and the Variable region remain to be investigated. The different non-coding sequences, especially the Variable region, might be important for the difference in the biological activities of the CTX phages; for example, the replication efficiency of the CTX phage and directionality of the integration into chromosome 1 or chromosome 2 of *V. cholerae*.

The replication of CTX-cla and CTX-2 has not been demonstrated, while the replication of CTX-1, CTX-O139, and RS1 element has been well-described [6, 9, 20]. The Variable regions in CTX-cla and CTX-O139 are homologous, but the non-coding sequence of CTX-O139 contains an additional 194 nucleotides. The replication efficiencies of CTX-cla and CTX-O139 might be altered by the Variable region and the additional sequence.

In addition to the efficiency of the replication of CTX phages, the chromosomal integration of CTX phage might be governed by the non-coding sequence. CTX-cla or CTX-2 have been found to be integrated on chromosomes 1 and 2, whereas CTX-1, CTX-O139, and RS1 have been identified only in chromosome 1 [13]. The exact role of the non-coding sequence in the replication and integration of CTX phage could be elucidated if the replication of CTX-cla and CTX-2 is established.

## Acknowledgments

This work was supported by the research fund of Hanyang University (HY-2012-P).

## References

1. Das B, Bischerour J, Barre FX. 2011. VGJΦ integration and excision mechanisms contribute to the genetic diversity of *Vibrio cholerae* epidemic strains. *Proc. Natl. Acad. Sci. USA* **108**: 2516-2521.
2. Das B, Bischerour J, Val ME, Barre FX. 2010. Molecular keys of the tropism of integration of the cholera toxin phage. *Proc. Natl. Acad. Sci. USA* **107**: 4377-4382.
3. Das B, Martinez E, Midonet C, Barre FX. 2013. Integrative mobile elements exploiting Xer recombination. *Trends Microbiol.* **21**: 23-30.
4. Davis BM, Kimsey HH, Chang W, Waldor MK. 1999. The *Vibrio cholerae* O139 Calcutta bacteriophage CTXΦ is infectious and encodes a novel repressor. *J. Bacteriol.* **181**: 6779-6787.
5. Davis BM, Moyer KE, Boyd EF, Waldor MK. 2000. CTX prophages in classical biotype *Vibrio cholerae*: functional phage genes but dysfunctional phage genomes. *J. Bacteriol.* **182**: 6992-6998.
6. Davis BM, Waldor MK. 2000. CTXΦ contains a hybrid genome derived from tandemly integrated elements. *Proc. Natl. Acad. Sci. USA* **97**: 8572-8577.
7. Faruque SM, Tam VC, Chowdhury N, Diraphat P, Dziejman M, Heidelberg JF, *et al.* 2007. Genomic analysis of the Mozambique strain of *Vibrio cholerae* O1 reveals the origin of El Tor strains carrying classical CTX prophage. *Proc. Natl. Acad. Sci. USA* **104**: 5151-5156.
8. Heidelberg JF, Eisen JA, Nelson WC, Clayton RA, Gwinn ML, Dodson RJ, *et al.* 2000. DNA sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*. *Nature* **406**: 477-483.
9. Kamruzzaman M, Robins WP, Bari SMN, Nahar S, Mekalanos JJ, Faruque SM. 2014. RS1 satellite phage promotes diversity of toxigenic *Vibrio cholerae* by driving CTX prophage loss and elimination of lysogenic immunity. *Infect. Immun.* **82**: 3636-3643.
10. Kaper JB, Morris JG Jr, Levine MM. 1995. Cholera. *Clin. Microbiol. Rev.* **8**: 48-86.
11. Kim EJ, Lee CH, Nair GB, Kim DW. 2015. Whole-genome sequence comparisons reveal the evolution of *Vibrio cholerae* O1. *Trends Microbiol.* **23**: 479-489.
12. Kim EJ, Lee D, Moon SH, Lee CH, Kim SJ, Lee JH, *et al.* 2014. Molecular insights into the evolutionary pathway of *Vibrio cholerae* O1 atypical El Tor variants. *PLoS Pathog.* **10**: e1004384.
13. Lee JH, Choi SY, Jeon YS, Lee HR, Kim EJ, Nguyen BM, *et al.* 2009. Classification of hybrid and altered *Vibrio cholerae* strains by CTX prophage and RS1 element structure. *J. Microbiol.* **47**: 783-788.
14. Lee JH, Han KH, Choi SY, Lucas ME, Mondlane C, Ansaruzzaman M, *et al.* 2006. Multilocus sequence typing (MLST) analysis of *Vibrio cholerae* O1 El Tor isolates from Mozambique that harbour the classical CTX prophage. *J. Med. Microbiol.* **55**: 165-170.
15. McLeod SM, Kimsey HH, Davis BM, Waldor MK. 2005. CTXΦ and *Vibrio cholerae*: exploring a newly recognized type of phage-host cell relationship. *Mol. Microbiol.* **57**: 347-356.

16. Midonet C, Das B, Paly E, Barre F-X. 2014. XerD-mediated FtsK-independent integration of TLC $\Phi$  into the *Vibrio cholerae* genome. *Proc. Natl. Acad. Sci. USA* **111**: 16848-16853.
17. Mutreja A, Kim DW, Thomson NR, Connor TR, Lee JH, Kariuki S, *et al.* 2011. Evidence for several waves of global transmission in the seventh cholera pandemic. *Nature* **477**: 462-465.
18. Nguyen BM, Lee JH, Cuong NT, Choi SY, Hien NT, Anh DD, *et al.* 2009. Cholera outbreaks caused by an altered *Vibrio cholerae* O1 El Tor biotype strain producing classical cholera toxin B in Vietnam in 2007 to 2008. *J. Clin. Microbiol.* **47**: 1568-1571.
19. Safa A, Nair GB, Kong RY. 2010. Evolution of new variants of *Vibrio cholerae* O1. *Trends Microbiol.* **18**: 46-54.
20. Waldor MK, Mekalanos JJ. 1996. Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272**: 1910-1914.