**Review Article**


**Molecular mechanism of acquisition of the cholera toxin genes**

Bhabatosh Das, Julien Bischerour & François-Xavier Barre

*CNRS, Centre de Génétique Moléculaire, Gif-sur-Yvette & Université Paris-Sud, Orsay, France*

Received June 14, 2010

One of the major pathogenic determinants of *Vibrio cholerae*, the cholera toxin, is encoded in the genome of a filamentous phage, CTXΦ. CTXΦ makes use of the chromosome dimer resolution system of *V. cholerae* to integrate its single stranded genome into one, the other, or both *V. cholerae* chromosomes. Here, we review current knowledge about this smart integration process.

**Key words** dif - site-specific recombination - XerC - XerD

**Introduction**

Most bacteriophages are detrimental to their host metabolism. However, phages also participate in the horizontal transfer of genes among bacteria because their genome can harbour other genes than those strictly required for their life cycle. This can be highly beneficial to the bacterial host. Indeed, many bacterial virulence factors are associated with phage-like DNA sequences. More strikingly, the exotoxins produced by many pathogenic bacteria are encoded in the genome of lysogenic phages. This is notably the case in *Bordetella avium*, *Clostridium botulinum*, *Corynebacterium diphtheriae*, *Escherichia coli*, *Pseudomonas aeruginosa*, *Shigella dysenteriae*, *Staphylococcus aureus* and *Streptococcus pyogenes*. The integrated prophages harboured by these bacteria profit from the multiplication of their host in the environment, which is in turn favoured by the virulence factors they bring to their host.

The study of *Vibrio cholerae*, the agent of the deadly diarrhoeal disease cholera, provides a fascinating case of such a bacterium-phage co-evolution. *V. cholerae* is the host for a variety of phages, commonly known as vibriophages, which can be lytic, non-lytic, virulent or temperate. On the one hand, phage predation of *V. cholerae* has been reported to be a factor that influences seasonal epidemics of cholera. On the other hand, one of the major virulence factors of *V. cholerae*, cholera toxin, is encoded in the genome of an integrated prophage CTXΦ. Furthermore, different variants of the phage CTXΦ exist, which participate in the genetic diversity of epidemic causing cholera strains. Two different attachment sites were found for this family of phages on the *V. cholerae* genome. They correspond to the dimer resolution sites of the two *V. cholerae* chromosomes, dif1 and dif2. Indeed, in contrast to most other lysogenic phages, such as bacteriophage λ, CTXΦ does not encode its integrase, but makes use of XerC and XerD, the two host-encoded tyrosine recombinases that normally function to resolve chromosome dimers. This mode of integration is all the more intriguing since CTXΦ phages belong to the filamentous phage family, which are generally
not lysogenic and which harbour a single stranded circular genome. Nevertheless, CTXΦ-like prophages were found integrated in the genome of several bacterial species, notably in pathogenic *E. coli* strains and in *Yersinia pestis*. Finally, it is remarkable to observe that many filamentous phages and/or genetic elements other than CTXΦ seem to have hijacked the chromosome dimer resolution system of *V. cholerae* for integration. Thus, TLC21, VEJ22, VGJ23, VSK24, VSKK (AF452449), KSF-1Φ24, fs125, fs226, f23724, were all found to be integrated at dif1 and/or dif2. Such a diversity of elements has not been observed in any other genera than the vibrios. Together, these elements participate in the dissemination of virulence factors among *V. cholerae* strains1,28,29 and in the emergence of new genetic variants of epidemic strains of *V. cholerae*. We review current knowledge on the integration mechanism of filamentous vibriophages that hijack the XerCD recombines, with a special focus on CTXΦ.

**CTXΦ integration model: exception or new paradigm?**

CTXΦ has a ~7-kb ss(+)DNA genome arranged in two modular structures, the “RS” and “core”. The core region harbours seven genes, which are *psh*, *cep*, *gliCTX*, *ace*, *zot*, *ctxA* and *ctxB*. While the *psh*, *cep*, *gliCTX*, *ace* and *zot* encoded proteins are needed for phage morphogenesis, the products of the *ctxAB* genes are not strictly required for the life cycle of the phage but are responsible for the severe diarrhoea associated with cholera. Three proteins, designated as RstR, RstA and RstB, are encoded in RS. Genetic analyses indicated that RstA is essential for phage replication and that RstB plays a crucial role in integration20. RstR acts as a transcriptional repressor by inhibiting the activity of PrstA, the only phage promoter required for CTXΦ replication and integration20. Several CTXΦ have been reported. These can be classified into four families based on the sequence of their *rstR* gene. These categories were designated as CTXΦET, CTXΦCl, CTXΦEc and CTXΦEv according to the host cells in which they were originally isolated31-33.

As mentioned earlier, the integration of CTXΦ into the *V. cholerae* genome depends on two host encoded tyrosine recombinases, XerC and XerD. XerC and XerD normally serve to resolve circular bacterial chromosome dimers generated by RecA mediated homologous recombination by adding a crossover at a specific 28 bp site *dif* on the chromosome16. The *dif* sites consist of specific 11-bp binding sites for each of the two Xer recombinases, separated by a 6-bp central region. These are generally located opposite to the origin of replication of bacterial chromosomes. Two *dif* sites are present on the genome of *V. cholerae*, one for each of the two circular chromosomes of the bacterium. Three different chromosome dimer resolution sites (*dif1, dif2* and *difG*) have been identified among the different *V. cholerae* strains characterized to date (Table I).

The ssDNA (+) genome of CTXΦ harbours two *dif* like sites (*attP1* and *attP2*). These are arranged in opposite orientation and are separated by ~90-bp DNA segment in the phage genome. Integration of CTXΦ at the *dif* loci of *V. cholerae* depends on the formation of a forked hairpin structure of 150 bp in the region encompassing *attP1* and *attP2* in the (+) ssDNA genome (Fig.1). The hybridization of *attP1* and *attP2* at the stem of this hairpin unmasks the phage attachment site, *attP(+)*. Integration occurs, XerC and XerD recombine this site with one of the two dimer resolution sites harboured by the host cell. This process only requires the catalytic activity of XerC: a single pair of strands is exchanged, which results in the formation of a pseudo-Holliday junction.

A proof of principle for this mechanism of integration was originally obtained for the El Tor variant of CTXΦ and *dif1* based on *in vivo* work performed in *Escherichia coli* and *in vitro* work performed with the *E. coli* Xer recombinases. Later on, a sensitive and quantitative assay was developed to confirm the ssDNA(+) integration model of CTXΦET into the *dif1* site of a *V. cholerae* El Tor strain36. This system was also used to define rules of compatibilities between the phage attachment sites harboured by the different CTXΦ variants characterized to date and their host dimer resolution sites: integration is solely determined by possibility to form Watson-Crick or Wobble base pair interactions to stabilize the exchange of strands promoted by XerC-catalysis between the phage attachment site and its target dimer resolution site (Table II and Fig. 1). These rules explain how integration of CTXΦET is restricted to *dif1*, how CTXΦCl can target both *dif1* and *dif2*, and how a third CTXΦ variant

<table>
<thead>
<tr>
<th>Site</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>dif1</em></td>
<td>AGTGCATTATTTGATGTTAGTAAAT</td>
</tr>
<tr>
<td><em>dif2</em></td>
<td>AATGCATTATTTGATGTTAGTAAAT</td>
</tr>
<tr>
<td><em>difG</em></td>
<td>AGTGCATTATTTGATGTTAGTAAAT</td>
</tr>
</tbody>
</table>

Source: Ref. 36
targets \textit{difG} (Table II). This single stranded integration model is not restricted to CTX\(\Phi\). Analysis of the \textit{attP} sites of CUS-1\(\Phi\) and Ypf-\(\Phi\) phages revealed features for direct ssDNA integration into the chromosome dimer resolution site harboured by their respective host cells\(^{18}\). Another family of mobile genetic element, the integrons, also integrates in the bacterial chromosome via a single stranded intermediate\(^{39}\).

### Integration mechanism of CTX\(\Phi\)-associated genetics elements

Several filamentous phages other than CTX\(\Phi\) are found to be integrated at the \textit{dif} loci of \textit{V. cholerae}\(^{13,22,23}\). To date, there is no report about their particular integration mechanism. Like CTX\(\Phi\), they do not encode a dedicated recombinase. In addition, a 29-bp \textit{dif} like sequence can be identified in many of them

---

**Table II.** Sequences of the \textit{dif}-like sites harboured by CTX\(\Phi\) variant

<table>
<thead>
<tr>
<th>CTX(\Phi) variant</th>
<th>\textit{attP} sequence</th>
<th>Integration site</th>
<th>Accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td>El Tor</td>
<td>AGTGCCTATTAGTGCGCCGCCGCA TTATGTTGAGG (\textit{attP}1)</td>
<td>\textit{dif}\textsubscript{1}</td>
<td>VCU83796</td>
</tr>
<tr>
<td></td>
<td>AATGCGTATTAGTGCCCTTAATGTTTACGG (\textit{attP}2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classical</td>
<td>AGTGCCTATTAGTGCGCCGCCGCA TTATGTTGAGG (\textit{attP}1)</td>
<td>\textit{dif}\textsubscript{1}</td>
<td>AY349175</td>
</tr>
<tr>
<td></td>
<td>AATGCGTATTAGTGCCCTTAATGTTTACGG (\textit{attP}2)</td>
<td>\textit{dif}\textsubscript{2}</td>
<td></td>
</tr>
<tr>
<td>Calcutta</td>
<td>AGTGCCTATTAGTGCGCCGCCGCA TTATGTTGAGG (\textit{attP}1)</td>
<td>\textit{dif}\textsubscript{1}</td>
<td>AF110029</td>
</tr>
<tr>
<td></td>
<td>AATGCGTATTAGTGCCCTTAATGTTTACGG (\textit{attP}2)</td>
<td>\textit{dif}\textsubscript{2}</td>
<td></td>
</tr>
<tr>
<td>G</td>
<td>AGTGCCTATTAGTGCGCCGCCGCA TTATGTTGAGG (\textit{attP}1)</td>
<td>\textit{dif}\textsubscript{G}</td>
<td>AF416590</td>
</tr>
<tr>
<td></td>
<td>AATGCGTATTAGTGCGCCGCCGCA TTATGTTTACGG (\textit{attP}2)</td>
<td>\textit{dif}\textsubscript{G}</td>
<td></td>
</tr>
</tbody>
</table>

\textit{Source:} Ref. 40
Table III. Sequences of the dif-like sites harboured by other vibriophages

<table>
<thead>
<tr>
<th>Phage</th>
<th>Genome size (kb)</th>
<th>attP sequence</th>
<th>Host</th>
<th>Integration site</th>
<th>Accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td>VEJ</td>
<td>6.8</td>
<td>ACTTCGCATTA TGTCGGC TTATGGTAAAA</td>
<td>V. cholerae</td>
<td>dif1</td>
<td>NC012757</td>
</tr>
<tr>
<td>VGJ</td>
<td>7.5</td>
<td>ACTTCGCATTA TGTCGGC TTATGGTAAAA</td>
<td>V. cholerae</td>
<td>dif1</td>
<td>AY242528.1</td>
</tr>
<tr>
<td>VSK</td>
<td>6.9</td>
<td>ACTTCGCATTA TGTCGGC TTATGGTAAAA</td>
<td>V. cholerae</td>
<td>dif1</td>
<td>NC003327</td>
</tr>
<tr>
<td>VSKK</td>
<td>6.8</td>
<td>ACTTCGCATTA TGTCGGC TTATGGTAAAA</td>
<td>V. cholerae</td>
<td>dif1</td>
<td>AF452449</td>
</tr>
<tr>
<td>KSF1</td>
<td>7.1</td>
<td>UK</td>
<td>V. cholerae</td>
<td>UK</td>
<td>AY714348</td>
</tr>
<tr>
<td>fs1</td>
<td>6.3</td>
<td>UK</td>
<td>V. cholerae</td>
<td>UK</td>
<td>NC004306.1</td>
</tr>
<tr>
<td>fs2</td>
<td>8.6</td>
<td>AGTGCATCA TTATGGTAAAA</td>
<td>V. cholerae</td>
<td>dif1</td>
<td>AB002632</td>
</tr>
<tr>
<td>f237</td>
<td>8.7</td>
<td>AGTGCATCA TTATGGTAAAA</td>
<td>V. parahemolyticus</td>
<td>dif1</td>
<td>NC002362</td>
</tr>
</tbody>
</table>

UK, unknow; Source: Ref. 40

Fig. 2. Putative mechanism of lysogenic conversion by the second type of filamentous phages that are found integrated into the chromosomal dimer resolution sites of V. cholerae. It is, therefore, very likely that these phages take control of the host XerC and XerD recombinases to integrate into the genome of their host. However, the presence of a single putative XerCD binding site on their genome makes it unlikely that the ssDNA form of their genome is directly used as a substrate for integration. We rather favour a model in which the double stranded replicative form of these phages is used for integration (Fig. 2). We are currently investigating this model using the tools we have developed for the study of CTXΦ.

Interestingly, the two TLC elements integrated in strain N16961 are flanked by the half of the dif sequence (TGTCGGCATTA TGATAG for one and AGTGCATATTGA TGATAG for the other). It is, therefore, reasonable to argue that their integration might be linked to the activity of the Xer recombinases.

Future prospects

The particular mode of integration of CTXΦ raises several questions. First, the efficiency of integration of a circular single stranded DNA molecule harbouring the sole attachment site of CTXΦ is very low. However, it becomes extremely efficient when the RS region of the phage is included. One likely explanation is that constant production and/or stabilization of the phage single stranded circular genome compensate for the instability of single stranded DNA in bacterial cells. RstB, which has been shown to be a single stranded DNA binding protein, could play a role in the stabilization of the integration substrate. Accordingly, its biochemical properties and sequence differ from those of the single stranded DNA binding protein encoded in the genome of VGJΦ, a phage that seems to rely on double stranded DNA integration. Second, only one pair of strands is exchanged between the single stranded DNA genome of CTXΦ and the double stranded DNA genome of its host, which leaves open the question of how the resulting pseudo-Holliday junction intermediate is processed. Is it stably maintained until the next round of bacterial DNA replication or processed by some host DNA repair machinery? What occurs when the replication fork collides against this
unusual structure? Finally, it is intriguing that so many phages take advantage of the Xer recombination system of vibrios as compared to other bacterial species. We wonder if it could be related to the particular life style and environment of the vibrios and/or their particular genome structure and management.

References


