

## Effect of deletion mutation on the recombination activity of Cre recombinase

Liu Rongrong, Wang Lixia and Lin Zhongping<sup>✉</sup>

The National Laboratory of Protein Engineering and Plant Genetic Engineering, Peking University, Beijing P.R.China; <sup>✉</sup>e-mail: [linzp@pku.edu.cn](mailto:linzp@pku.edu.cn)

Received: 14 December, 2004; revised: 31 January, 2005; accepted: 02 February, 2005  
available on-line: 15 May, 2005

Cre recombinase from bacteriophage P1 is widely used in both *in vitro* and *in vivo* DNA manipulations. Based on a structural and functional analysis, three deleted *cre* mutants were constructed and expressed in *Escherichia coli*. Mutated recombinases were purified and their recombination activities were determined *in vitro*. Our results revealed that the mutant with amino-terminal deletion retains the recombination activity as high as wild type Cre; however, the carboxy-terminal deletion and the middle region deletion both lead to a complete loss of the recombinase function.

**Keywords:** Cre recombinase, deletion mutation, *in vitro* recombination assay

Cre is a site-specific recombinase from bacteriophage P1 with a full length of 343 amino-acid residues and a molecular mass of 38.5 kDa. It can cause recombination between two directly or inversely oriented 34 bp *loxP* sites on double-stranded DNA (Abremski & Hoess, 1985; Sternberg *et al.*, 1986). Crystal structure shows that Cre folds into two distinct domains, a smaller amino-terminal domain and a larger carboxy-terminal one, which are separated by a short linker. When interacting with DNA, these two domains form a C-shaped clamp around the *loxP* site and contact closely with both the major and minor grooves (Abremski & Hoess, 1985). The 34 bp *loxP* site has two 13 bp inverted repeats called recombinase binding elements (RBE), flanking an 8 bp asymmetric central strand exchange or crossover region that determines the direction of *loxP*. With each RBE binding to one Cre, one *loxP* site can associate with two Cre molecules. In the recombination reaction, four Cre subunits binding to two *loxP* sites combine to form a tetramer, which enables the two *loxP* sites to get close and exchange DNA strands.

In the past two decades, considerable studies have been held to illustrate Cre's structure-function correlation (VanDuyne, 2001; Guo *et al.*, 1997; Hoess *et al.*, 1990). Except for the first 20 amino-acid residues, which have not been resolved so far in

the published crystal structures, the amino-terminal domain is made up of five continual  $\alpha$ -helices, A, B, C, D, and E. Two neighboring Cre subunits bound to different DNA strands in the recombination tetramer interact by their helices A which extend away from the other helices. Helices B and D contact the major groove of the substrate DNA. Helix E associates with DNA backbone as well as its counterpart of another Cre subunit in the tetramer bound to the same DNA strand. The carboxy-terminal domain contains nine helices, F, G, H, I, J, K, L, M and N, most of which contact with the backbone or the major and minor grooves (Guo *et al.*, 1997). Arg173 on helix H, His289 and Arg292 on helix K and Tyr324 on helix M compose a characteristic "R-H-T-Y" structure found in all integrases, the most important family of site-specific recombinases (Abremski & Hoess, 1992). In Cre's tertiary structure, these four residues associate to form a conservative conformation where the Cre catalytic center is located. Helix N of the carboxy-terminal domain is apart from the rest of the molecule. It extends into a hydrophobic pocket of the neighboring subunit and interacts with it by hydrophobic interactions, so that four Cre subunits crosslink in a cyclic manner (Van Duyne, 2001).

Based on what has been learned about the relationship between Cre's structure and function, we

**Abbreviations:** IPTG, isopropyl- $\beta$ -D-thiogalactoside; mCreC, mCreM and mCreN, carboxy-terminal, middle region and amino-terminal deleted Cre recombinase mutant, respectively; RBE, recombinase binding element.

designed three deletion mutations. As the amino-acid residues at the amino terminus of the molecule are unresolved and suspected to form a random coil, we have deleted the first 12 residues and thus obtained an amino-terminal deleted Cre mutant (mCreN) to test whether this fragment is necessary for Cre's function. Helix N of the carboxy-terminal domain has been suggested to mediate interaction between Cre subunits. Petyuk *et al.* (2004) have inserted a pentapeptide in position 333 of Cre recombinase and got a derivative without cleavage competence, therefore, a carboxy-terminal deleted Cre mutant (mCreC) lacking the last 14 residues has been constructed to check the effect of its absence. Additionally, a middle region deleted Cre mutant (mCreM) lacking 18 amino-acid residues (Asp126–Asp143) was designed to examine whether the linker between the two domains is dispensable for Cre's recombination activity, as Lee and Sadowski (2003) have found out that single mutation of Glu129 or Gln133 abolish Cre's catalytic ability. All these mutations avoid Cre's catalytic site, main DNA contacting sites and areas with high sequence similarity with other site-specific recombinases: Leu203–Arg223, Ala285–Ser305 and Ile309–Asp329 (Wierzbicki *et al.*, 1987).

## MATERIALS AND METHODS

**Plasmids.** The plasmid harboring full-length *cre*, pMM23, was donated by Dr. David Ow of the United States Department of Agriculture. The 4.2 kb plasmid pGLGFP used as the substrate for *in vitro* recombination assay contains two directly oriented *loxP* sites flanking a 1.1 kb *gfp* expression unit. As illustrated in Fig. 1, pGLGFP can undergo an intramolecular recombination mediated by Cre.

**Deletion mutations.** Six PCR primers were synthesized according to *cre* sequence (Table 1). Using pMM23 as template, we amplified *cre* and its three deletion mutants with different primer pairs: p1 and p2 were used to amplify full-length *cre*; p3 and p2 were used to amplify *mCreN*; p1 and p4 were used to amplify *mCreC*; PCR product of p1 and p5 and PCR product of p2 and p6 were digested with *HindIII* and then ligated to form *mCreM*. The PCR products were sequenced before cloning into expression vector.

**Gene expression in *E. coli* BL21(DE3).** *Cre* and its mutants were cloned into expression vector pET-29a, transformed into *E. coli* strain BL21(DE3), and 1 mM IPTG was used to induce gene expression. After 2 h of induction, bacteria were collected by centrifugation and protein expression was examined by SDS/PAGE.

**Isolation, dissolution and renaturation of inclusion body.** The collected bacteria were suspended in STET buffer (0.1 M NaCl, 10 mM Tris/HCl, pH

8.0, 1 mM EDTA, pH 8.0, 5% Triton X-100), then disrupted with an ultrasonic disintegrator and centrifuged at 12000 r.p.m. for 15 min at 4°C to recover the inclusion body pellet. The pellet was washed with STET containing 1 M urea, centrifuged again, and dissolved with STET containing 8 M urea as denaturant. The dissolution process lasted overnight at 4°C. Afterwards, 9 volumes of renaturation buffer (50 mM  $\text{KH}_2\text{PO}_4$ , pH 10.7, 1 mM EDTA, pH 8.0, 50 mM NaCl, 2 M urea) was used to dilute the denaturant, pH was adjusted to 10.7 with KOH and the protein was kept overnight at 4°C for refolding. Finally, the pH was adjusted to 8.0 with HCl and the solution was centrifuged at 12000 r.p.m. for 15 min at 4°C after 30 min incubation at room temperature. The supernatant contains renatured proteins.

***In vitro* recombination assay.** Recombination was carried out in reaction buffer (0.5 M Tris/HCl, pH 7.5, 0.33 M NaCl, 0.05 M spermidine, 500 µg/ml BSA). Adequate pGLGFP DNA and purified recombinase were added into reaction buffer and incubated at 37°C for 30 min. The reaction was stopped by heating to 70°C for 10 min and the mixture was extracted with chloroform/isoamyl alcohol (24:1, v/v). Finally, DNA was precipitated with ethanol, digested with the restriction endonuclease *ClaI* and run on a 1% agarose gel.

## RESULTS AND DISCUSSION

### Expression and purification of Cre and its mutants

After 2 h IPTG induction, Cre, mCreN and mCreC were expressed at a high level, while mCreM's expression level was much lower but detectable; their inclusion bodies were enriched in the recovered pellets. The inclusion bodies were dissolved in 8 M urea and the denaturant was diluted with renaturation buffer for the proteins to refold. The purity of the refolded recombinases was examined by SDS/PAGE (Fig. 2). The renatured recombinase with recombination activity could carry out *in vitro* recombination efficiently and consistently, which could exclude the possibility that a negative recombination result might be caused by incorrect protein folding.

### *In vitro* recombination assay

Intramolecular recombination between the two direct *loxP* sites will incise pGLGFP into two smaller circular DNA molecules (Fig. 1). It should be pointed out that the reaction Cre catalyzes is reversible and reaches a dynamic equilibrium, therefore when recombination is finished, the reaction system will contain certain amount of intact pGLGFP besides the two recombination products. It has been

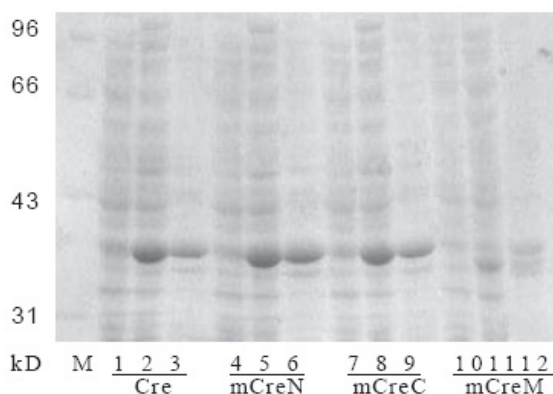
**Table 1. Sequences of PCR primers used to create Cre deletion mutation.**

Bold letters show restriction endonuclease sites.

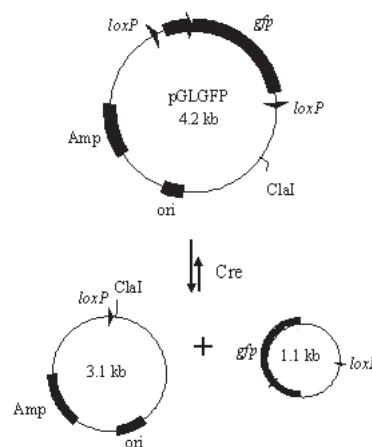
Primer	Sequence (5'→3')	Restriction site
p1	CGATATAC <b>CATATG</b> TCCAATTTACTGACCGTACAC	<i>NdeI</i>
p2	TCG <b>AGATCT</b> TGCTAATCGCCATCTTCCAG	<i>BglIII</i>
p3	GACTG <b>CATATG</b> TTACCGGTCGATGCAACGAGTG	<i>NdeI</i>
p4	TCG <b>AGATCT</b> TTAACTATCCAGGTTACGGAT	<i>BglIII</i>
p5	CGC <b>AGCTT</b> TTTCCTTTTCGGATCCGCCGC	<i>HindIII</i>
p6	CGC <b>AGCTT</b> CGTTCACTCATGGAAATAGC	<i>HindIII</i>

noticed that neither increasing recombinase concentration nor lengthening reaction time could evidently enhance the recombination efficiency. When DNA in the reaction mixtures is digested with *ClaI*, both intact pGLGFP and the larger recombination product will be linearized. Thus, after recombination and *ClaI* digestion, three bands can be seen on agarose gel: a 4.2 kb band of linearized pGLGFP, a 3.1 kb band of linearized larger recombination product, and another band of the smaller circular DNA molecule yielded by recombination. Our results showed that mCreN retains the recombination activity nearly as high as wild type Cre, while both mCreC and mCreM have almost completely lost their catalytic ability (Fig. 3).

The first 20 amino-acid residues at the amino terminus ahead of helix A have not been resolved so far in the published literature and are thus considered as unstructured. Lacking the first 12 residues of wild type Cre, mCreN retains the recombination activity, which indicates that this fragment does not play an important role in Cre functions, including binding to *loxP* site, interacting with other subunits

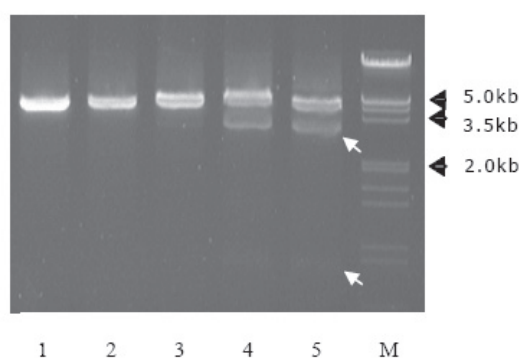
**Figure 2. Expression and purification of Cre recombinase and its mutants.**

Lanes 1, 4, 7 and 10: total bacterial protein before induction; lanes 2, 5, 8 and 11: total bacterial protein after 2 h IPTG induction; lanes 3, 6, 9 and 12: partially purified and refolded recombinases; M: protein molecular mass markers.

**Figure 1. Intramolecular recombination of pGLGFP catalyzed by Cre.**

Intramolecular recombination between two directly oriented *loxP* sites of the 4.2 kb plasmid pGLGFP will yield two smaller plasmids of 3.1 kb and 1.1 kb. The reaction is reversible.

and cleaving the substrate. The Cre catalytic site and main DNA-binding sites are entirely within the amino-terminal and the carboxy-terminal domains, while the function of the linker between the two domains is unclear. In our experiments, the Cre mutant deprived of 18 amino-acid residues, Asp126–Asp143, has completely lost its recombination activity. This may be due to a conformation rigidity change. Besides, this may also be because the two domains in the mutant are spatially too close to form a C-shaped

**Figure 3. In vitro recombination assay of Cre and its mutants.**

Lane 1, pGLGFP digested with *ClaI* as a negative control; 2–5, pGLGFP DNA incubated with purified mCreC, mCreM, mCreN, and Cre, respectively, followed by *ClaI* digestion. Notice that mCreN kept the *in vitro* recombination activity as the wild type Cre, with white arrow heads indicating out recombination products; however, mCreC and mCreM both have almost entirely lost their function, with the gel revealing only one band of pGLGFP DNA. M, DNA molecular size markers.

clamp, which in turn makes the recombinase unable to bind to DNA double strands from opposite sides. A recent study indicated that the linker might have a more complicated function in Cre activity: Glu129 can interact with Val85 of the same subunit and Lys201 of the neighboring subunit in the tetramer; Gln133 is spatially near to helix M, where the catalytic Tyr324 lies, and can also interact with the nucleotides near the DNA nick produced by the recombinase during recombination (Lee & Sadowski, 2003). These results suggest that the linker between the two domains is not completely flexible and does not merely function as a joining. It might determine the relative position of the two domains *via* interacting with certain amino-acid residues within both of them. Besides, the linker might also participate in maintaining the conformation of the tetramer and the recombinase–DNA complex. What should also be taken into consideration is that the deletion of mCreM includes not only the linker but also a part of helix F which contains a conservative Glu, and this might also contribute to its loss of function.

In Cre tetramer helix N of each subunit extends into its neighboring subunit and associates with it by hydrophobic interaction, through which four Cre molecules crosslink into a circle (VanDuyne, 2001). Therefore, helix N is essential for the tetramer's structure, and once this function is destroyed, Cre subunits may not associate into a correct quaternary structure and recombination cannot occur. Helix N is absent in mCreC, which, consistent with our anticipation, results in the loss of recombinase activity. Additionally, in the 329 aa mCreC, catalytic Tyr324 is very near to the carboxy terminus. This is very likely to cause an invalid conformation in the catalytic center and therefore the loss of recombination activity.

In one word, our work suggests that the first 12 amino-acid residues at the amino terminus of Cre recombinase do not play an important role in the re-

combination reaction; however, the linker between the two protein domains and the last 14 amino-acid residues at the carboxy terminus are both indispensable for Cre activity.

As Cre recombinase functions in a simple, effective and consistent manner, it gradually becomes a widely used tool in changing genome structure and modifying gene expression (Sauer, 1993; Kilby *et al.*, 1993; Chen *et al.*, 2003). What's more, modified Cre recombinase has already been used to mediate rearrangement of DNA sequences flanked by *loxP* sites (Koresawa *et al.*, 2000). This study helps us to learn better the relationship between Cre structure and function, and offers useful information to those attempting to modify Cre for more effective and more convenient application in DNA manipulations.

## REFERENCES

- Abremski K, Hoess R (1985) *J Mol Biol* **181**: 351–362.  
Abremski K, Hoess R (1992) *Protein Eng* **5**: 87–91.  
Chen M, Wang LX, Peng XL, Xu HJ, Lin ZP (2003) *Acta Botanica Sinica* **45**: 1481–1488.  
Guo F, Gopaul DN, Van Duyne GD (1997) *Nature* **389**: 40–46.  
Hoess R, Abremski K, Irwin S, Kendall M, Mack A (1990) *J Mol Biol* **216**: 873–882.  
Kilby NJ, Snaith MR, Murray JA. (1993) *Trends Genet* **9**: 413–421  
Koresawa Y, Miyagawa S, Ikawa M, Matsunami K, Yamada M, Shirakura R, Okaba M (2000) *J Biochem* **127**: 367–372.  
Lee L, Sadowski PD (2003) *J Biol Chem* **278**: 36905–36915.  
Petyuk V, McDermott J, Cook M, Sauer B (2004) *J Biol Chem* **279**: 37040–37048.  
Sauer B (1993) *Methods Enzymol* **225**: 890–900.  
Sternberg N, Sauer B, Hoess R, Abremski K (1986) *J Mol Biol* **187**: 197–212.  
Van Duyne GD (2001) *Annu Rev Biophys Biomol Struct* **30**: 87–104.  
Wierzbicki A, Kendall M, Abremski K, Hoess R (1987) *J Mol Biol* **195**: 785–794.