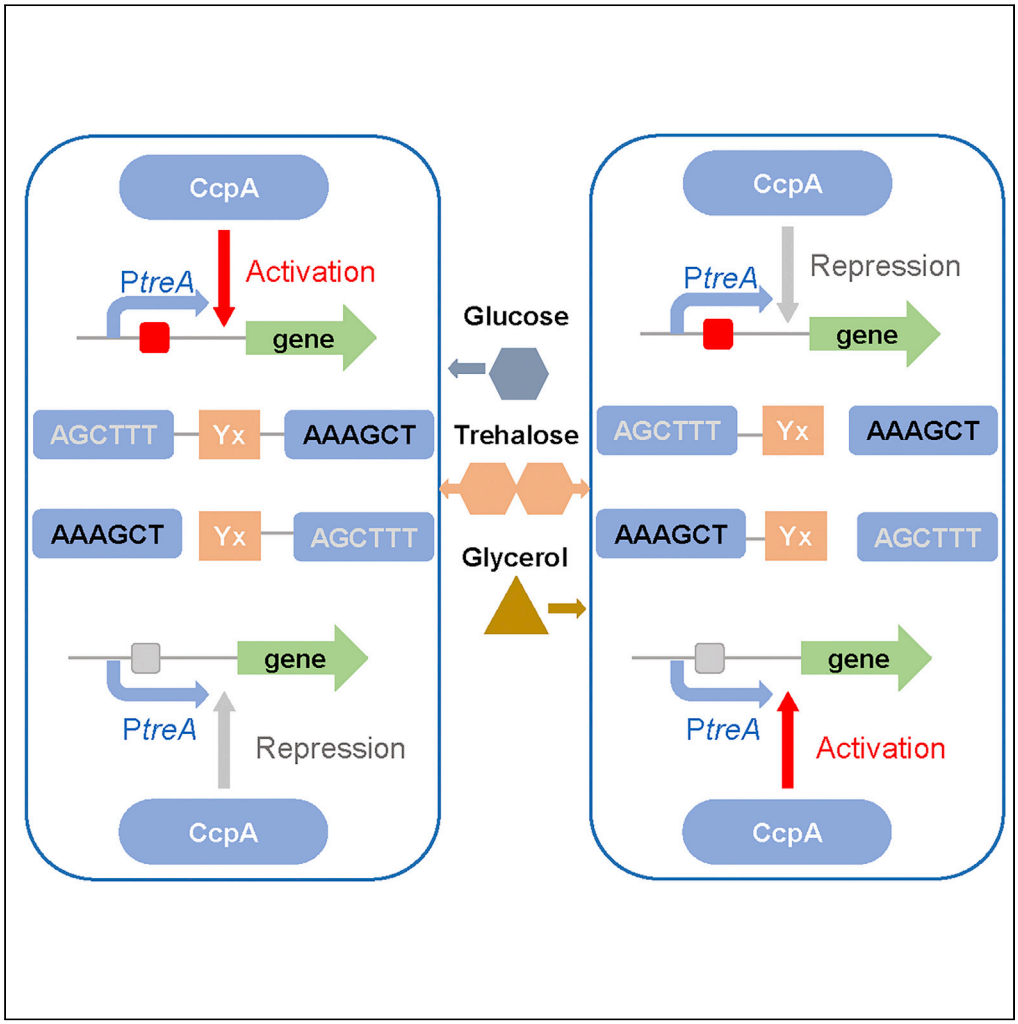
iScience

CellPress
OPEN ACCESS

Article

A new CcpA binding site plays a bidirectional role in carbon catabolism in *Bacillus licheniformis*



Fengxu Xiao, Youran Li, Yupeng Zhang, ..., Zhenghua Gu, Sha Xu, Guiyang Shi

liyouran@jiangnan.edu.cn (Y.L.)
gyshi@jiangnan.edu.cn (G.S.)

Highlights

A novel CcpA binding site CRE_{Tre} was identified and characterized

CRE_{Tre} consists of two palindrome arms of 6 nucleotides (AGCTTT/AAAGCT)

CRE_{Tre} is involved in a bidirectional regulation under the glucose stress

Xiao et al., iScience *24*, 102400
May 21, 2021
https://doi.org/10.1016/j.isci.2021.102400

iScience

CellPress
OPEN ACCESS

Article

A new CcpA binding site plays a bidirectional role in carbon catabolism in *Bacillus licheniformis*

Fengxu Xiao,[1,2,3] Youran Li,[1,2,3,*] Yupeng Zhang,[1,2,3] Hanrong Wang,[1,2,3] Liang Zhang,[1,2,3] Zhongyang Ding,[1,2,3] Zhenghua Gu,[1,2,3] Sha Xu,[1,2,3] and Guiyang Shi[1,2,3,4,*]

SUMMARY

***Bacillus licheniformis* is widely used to produce various valuable products, such as food enzymes, industrial chemicals, and biocides. The carbon catabolite regulation process in the utilization of raw materials is crucial to maximizing the efficiency of this microbial cell factory. The current understanding of the molecular mechanism of this regulation is based on limited motif patterns in protein-DNA recognition, where the typical catabolite-responsive element (CRE) motif is "TGWNANCGNTNWCA". Here, CRE_{Tre} is identified and characterized as a new CRE. It consists of two palindrome arms of 6 nucleotides (AGCTTT/AAAGCT) and an intermediate spacer. CRE_{Tre} is involved in bidirectional regulation in a glucose stress environment. When AGCTTT appears in the 5′ end, the regulatory element exhibits a carbon catabolite activation effect, while AAAGCT in the 5′ end corresponds to carbon catabolite repression. Further investigation indicated a wide occurrence of CRE_{Tre} in the genome of *B. licheniformis*.**

INTRODUCTION

To cope with various environments, bacteria have developed a sophisticated carbon source utilization mechanism, which is mainly characterized by hierarchical uptake and metabolism of mixed carbon sources (Jörg and Hillen, 1999; Sonenshein, 2007). Usually, this mechanism avoids a simultaneous utilization of all available carbohydrates in order to save cell energy. Two regulatory phenotypes, carbon catabolite activation (CCA) and carbon catabolite repression (CCR), are widely observed in this bacterial process (Görke and Stülke, 2008). Previous studies found that carbon catabolite protein A (CcpA) played a vital role in the regulation of both catabolism and anabolism (Yoshida et al., 2001; Xiao and Xu, 2007). To carry out its regulation, CcpA or the complex CcpA-Hpr-Ser46-P (a combination of CcpA and the Ser46 phosphorylated form of a histidine-containing phosphocarrier protein [HPr]) needs to bind a cis-acting element—catabolite-responsive element (CRE)—either in the region upstream of the transcription start site or in the coding region (Schumacher et al., 2011). To date, a classic CRE motif "TGWNANCGNTNWCA" (where W stands for A or T and N for any base) (Weickert and Chambliss, 1990) has been identified and characterized. Researchers generally agreed that CRE sites were highly degenerated pseudo-palindromes and that the lack of stringent sequence conservation provides CcpA with a high degree of regulatory flexibility. Examples include "RRGAAAANGTTTTCWW" in *Clostridium difficile* (Antunes et al., 2012), TGTTATATAACA in *Clostridium acetobutylicum* (Willenborg et al., 2014), four CcpA binding sites (GAAGTTTTAAG; ATTTTTTGT; TATG AAAAATTTTAAAAAGTGGA; AGGCTTATCATAG) in *Streptococcus mutans* (Kim and Burne, 2017), and "TTTTYHWDHHWWTTTY" (Y represents base C or T, H represents base A, C, or T, W represents base A or T, and D represents base A, G, or T) in *Streptococcus suis* (Zhang et al., 2018). However, more differentiated CRE patterns are needed to explain how CcpA can bind such a diverse set of operator sites.

Bacillus licheniformis has been used widely as a microbial cell factory for enzymes and chemical production (Shi et al., 2019). It has exhibited obvious selective utilization of mixed carbon sources in industrial fermentation (Li et al., 2019). Various oligosaccharides, in particular, significantly affect cell growth and metabolism (Patricia et al., 2014). Much has been explained about the central role of CcpA in the coordinated regulation of catabolism and anabolism using the classic CRE motif to ensure optimal cell growth under varying environmental conditions (Fujita, 2014). However, a large number of CcpA target genes, revealed by microarray or RNA sequencing, do not contain the classical CRE motif (Ruud et al.,

[1]Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi 214122, People's Republic of China

[2]National Engineering Laboratory for Cereal Fermentation Technology, Jiangnan University, 1800 Lihu Avenue, Wuxi, Jiangsu Province 214122, People's Republic of China

[3]Jiangsu Provincial Research Center for Bioactive Product Processing Technology, Jiangnan University, Wuxi, China

[4]Lead contact

*Correspondence: liyouran@jiangnan.edu.cn (Y.L.), gyshi@jiangnan.edu.cn (G.S.)

https://doi.org/10.1016/j.isci.2021.102400

2015; Lin et al., 2013b). In this context, the specific molecular mechanism of CcpA target gene recognition remains elusive.

Another problem that must be considered is that CcpA is involved in both CCR and CCA, even under the same stress from glucose or other oligosaccharides. For example, some *Bacillus* exhibited a preference for malic acid over glucose (Bruckner and Titgemeyer, 2002; Asai et al., 2000). Transcriptome information showed that genes of the malic acid catabolism pathway were activated in the presence of this organic acid, while those within the glucose catabolism pathway were repressed (Kleijn et al., 2010). Glucose can also suppress the catabolism of other phosphotransferase system (PTS) carbohydrates, such as xylose and mannitol (Li et al., 2018; Xiao et al., 2020). The pleiotropic role of CcpA is related in part to its protein structure (Loll et al., 2007), while the target DNA site is also an important basis of its function. The current research on CcpA focuses on its CCR effect, the molecular mechanism of which has been explained using the classical CRE motif. However, little is known about the CCA mechanism and the related target DNA sites.

Trehalose is a disaccharide consisting of two glucose moieties. It also acts as a stress protectant in various bacteria (Okabe et al., 2020). The trehalose operon is reported to regulate the transportation of trehalose and the interconversion between this disaccharide and glucose (Eastmond and Graham, 2003); thus, in the past decade, it has attracted increasing research interest on the adaptability of microorganisms. Previous results proved that the trehalose operon could interact with CcpA in response to glucose or trehalose, where a classical CRE motif was identified in the trehalose promoter (Gosseringer et al., 1997). In this study, the structure of the trehalose operon of *B. licheniformis* was elucidated. A new CcpA binding site in the *treR* gene, CRE_{Tre}, was identified and characterized as a new CRE. The new motif presents completely different structural characteristics than those of the classical motif. In this study, the relationship between its structure and the CcpA-mediated CCR/CCA effect was investigated. The results provide a novel model for the regulation of the CcpA protein in *B. licheniformis*.

RESULTS

Discovery of a new CcpA protein binding site in the *B. licheniformis* trehalose operon

The trehalose operon consists of a promoter and three structural genes, *treP* (encoding $EIIBC^{Tre}$), *treA* (encoding trehalose 6-phosphate hydrolase), and *treR* (encoding the regulatory protein) (Schöck and Dahl, 1996). Two transcription elements (the P*treA* promoter and the repressor protein TreR) were identified in the trehalose operon. P*treA* is responsible for the transcription of the structural genes, *treA* and *treP*, and the TreR protein is capable of turning the promoter "on" or "off". Several potential CRE sites were located by searching using a typical CRE motif in P*treA* and *treR*. A sequence that is highly similar to the typical CRE motif (with only base pair 13 inconsistent), "TGAAAGCGCTATAA" (cre1) was found downstream of the −35 region in P*treA*. Next, to identify other possible CcpA binding sites, 200-bp P*treA* was divided into two fragments, an 88-bp fragment in the 5′ end (A) and a 102-bp fragment in the 3′ end (B) (Figure 1A). Additionally, the cre1 was deleted from fragment B, resulting in a fragment C. The above three fragments were subjected to electrophoretic mobility shift assay (EMSA) to investigate their affinity with CcpA. The results are shown in Figure 1B. CcpA did not recognize or bind to fragments A and C. However, it clearly showed that the shift of fragment B was blocked. This result indicated that there is only one CcpA binding site in P*treA*. Notably, Hpr (Ser-P) binds to CcpA, thereby inducing the binding of the complex to the cre site (Warner and Lolkema, 2003). Hence, the function of Hpr (Ser-P) was tested for CcpA binding *in vitro*. As shown in Figure S1, there were the same results for CcpA binding in the lane regardless of whether Hpr (Ser-P) was added. Hence, in the next experiment, the Hpr (Ser-P) was not added in the lane for the EMSA experiment.

Then, the sequence of *treR* was also searched for the typical CRE site. A sequence "AGACACCGCTTGGA" (cre2) was found between 664 and 677 bp from the initial codon. It had two bases (1 and 13) that did not fit with the typical CRE motif. Next, to identify other possible CcpA binding sites, *treR* was divided into four fragments, a 199-bp fragment in the 5′ end (D), a 202-bp fragment in the 5′ end (E), a 196-bp fragment in the 3′ end (F), and a 194-bp fragment (G) in the 3′ end (Figure 1C). The results are shown in Figure 1D. A DNA band shift was observed for fragments D, E, and G, suggesting that there exist CcpA binding sites in these fragments. Hence, it is necessary to verify the specific CcpA binding site in these fragments. The cre2 was first deleted from fragment G, resulting in a G-Δcre2 fragment. The EMSA result showed no affinity between this fragment and CcpA, indicating that cre2 was the sole binding site. After that, fragment D was

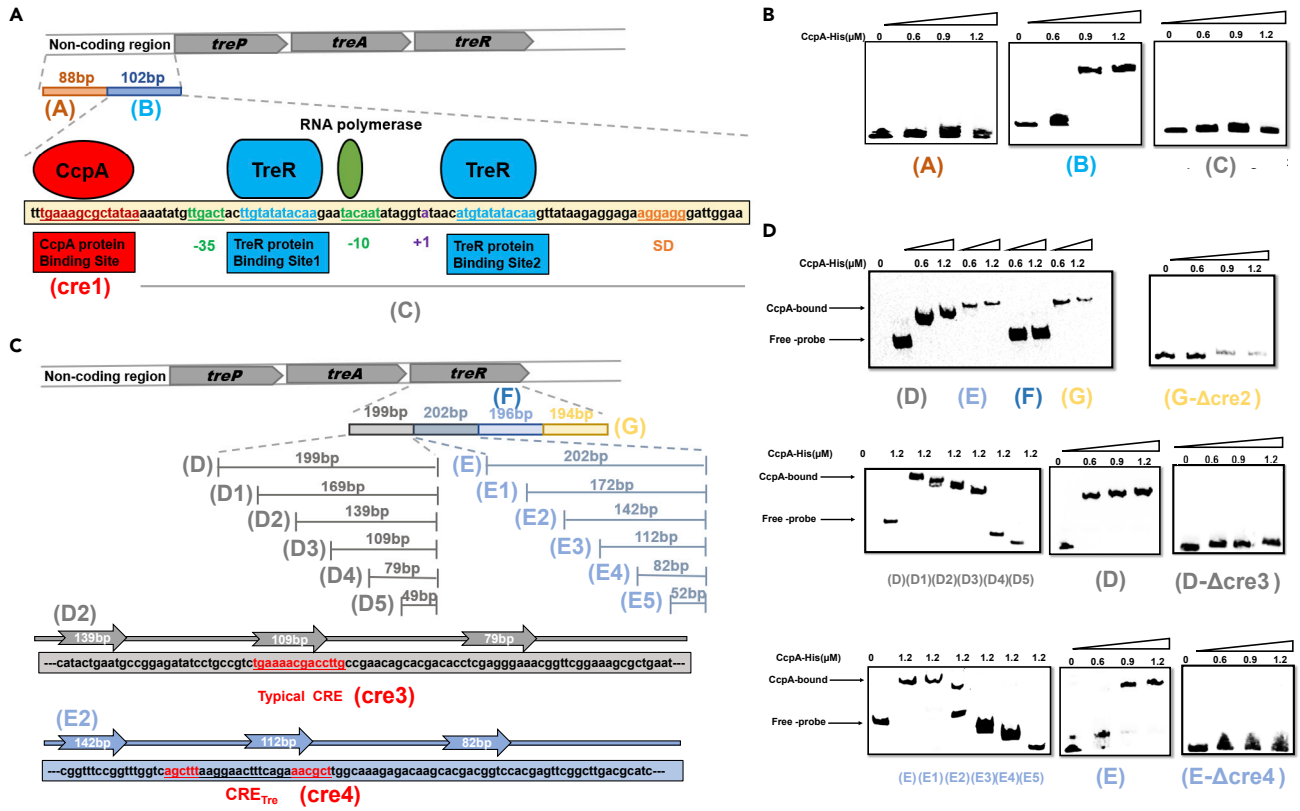


Figure 1. Identification of CcpA binding sites in trehalose operon

(A) The non-coding region of the trehalose operon was divided into two fragments ([fragment A] and [fragment B]). The CcpA binding region is shown in red, and the TreR binding region is shown in blue. The −10 region and −35 region are shown in green. The transcription start site is shown in purple. The SD sequence is shown in orange.

(B) EMSA of CcpA protein binding to three fragments (fragment A, fragment B, and fragment C) labeled with 5′ biotin.

(C) The region of *treR* gene was divided into four fragments (fragment D, fragment E, fragment F, and fragment G). The fragment D of the *treR* gene was further divided into six fragments (fragment D, fragment D1, fragment D2, fragment D3, fragment D4, and fragment D5). The fragment E of the *treR* gene was further divided into six fragments (fragment E, fragment E1, fragment E2, fragment E3, fragment E4, and fragment E5). The CcpA binding site in fragment D or fragment E is shown in red.

(D) EMSA of CcpA protein binding to four fragments (fragment D, fragment E, fragment F, and fragment G) labeled with 5′ biotin, six fragments labeled with 5′ biotin (fragment D, fragment D1, fragment D2, fragment D3, fragment D4, and fragment D5) that were derived from fragment D, and six fragments labeled with 5′ biotin (fragment E, fragment E1, fragment E2, fragment E3, fragment E4, and fragment E5) that were derived from fragment E.

gradually truncated at intervals of 30-bp bases from the 3′ end, thus resulting in the following fragments: 169-bp fragment (D1), 139-bp fragment (D2), 109-bp fragment (D3), 79-bp fragment (D4), and 49-bp fragment (D5) (Figure 1D). A significantly shifted band was found in fragments D1, D2, and D3. In contrast, the band was completely eliminated when fragment D4 or D5 was used, indicating that the binding site was located between fragments D4 and D5. The truncated 30 base fragment was scanned for possible CRE structure, and a sequence "TGAAAACGACCTTG" (cre3) was found. In the conservative region, there were three bases (10, 13,14) that did not fit with the typical CRE motif. Then, the cre3 was deleted from fragment D, resulting in a D-Δcre3 fragment. As predicted, the EMSA results showed that CcpA could no longer bind to fragment D-Δcre3 (Figure 1D).

Next, fragment E, another CcpA-binding fragment, was scanned for possible CRE structure. However, no classical CRE site was found. Then, the same method was carried out as described above, generating the following fragments: 172-bp fragment (E1), 142-bp fragment (E2), 112-bp fragment (E3), 82-bp fragment (E4), and 52-bp fragment (E5) (Figure 1C). The EMSA results showed that the CcpA protein lost its affinity for fragment E3, while it maintained its affinity for fragment E2, indicating that the binding site was located between these two fragments. Additionally, a 26-bp pseudopalindrome sequence (AGCTTT-aaggaactttc aga-AACGCT) (cre4) was found in this region. To validate its function, the cre4 was deleted from fragment

E, resulting in an E-Δcre4 fragment. The EMSA results showed no affinity between CcpA and fragment E-Δcre4, indicating that the pseudo-palindromic sequence (cre4) may be the binding site for the CcpA protein. Based on fragment E, two fragments, fragments E6 and E7, were artificially generated. These fragments contain cre4-1 (AGCGTT-aaggaactttcaga-AACGCT) and cre4-2 (AGCTTT-aaggaactttcaga-AAAGCT) (Figure S2). The results showed that fragment E7 has a significantly shifted band, while fragment E6 does not, indicating that cre4-2 is the novel CcpA binding site.

Taken together, the motif of the newly found CRE, CRE_{Tre}, is "AGCTTT-Yx-AAAGCT" (Y stands for any base, and X stands for the number of bases), which consists of a symmetrical region and intermediate spacer region. Compared with previously reported CRE sites in bacteria (*Bacillus*, *Lactobacillus*, and *Staphylococcus*), the novel CcpA binding site exhibits a high "AT" content feature.

Characterization of the CcpA binding motif CRE_{Tre}

CRE_{Tre} contains two parts, a 12-bp symmetrical region (AGCTTT/AAAGCT) and an intermediate spacer region. The effect of the nucleotide structure on the function remains elusive. Considering the nanoaffinity of CcpA for fragment A, the following fragments were constructed to determine the role of the two parts. First, the role of the intermediate spacer region was confirmed. To this end, three intermediate spacer fragments with different lengths, tre-26-1, tre-20-2, and tre-12-3, were constructed (Figure 2B). They were inserted into fragment A to obtain H1, H2, and H3, respectively (Figure 2A). All three constructed fragments showed shifted bands (Figure 2C), indicating that the truncated intermediate spacer region retained binding affinity with CcpA. Next, the role of the symmetrical region was confirmed. Three fragments, tre-26-4, tre-26-5, and tre-26-6, were constructed. These fragments contained artificially changed symmetrical regions, as shown in Figure 2B. They were also inserted into fragment A, yielding H4, H5, and H6, respectively. When the above three hybrid fragments were subjected to EMSA with CcpA, no shifted band was observed (Figure 2C). These results indicate that the symmetrical region is crucial for CcpA binding.

The above result indicates that the two inverted 6-bp repetitions are important for the binding of CcpA-CRE_{Tre}. Therefore, the question of whether each base is essential was explored. Every signal nucleotide within the 6-bp palindromes in fragment H1 was separately mutated, resulting in 12 derivative probes (T1, T2, T3, T4, T5, T6, T7, T8, T9, T10, T11, and T12) (Figure 3A). When 1.2 μM CcpA was used to incubate, the probes T1, T2, T3, T4, T5, T6, T7, T8, T9, and T10 exhibited high affinity for CcpA, while T11 and T12 showed significantly decreased affinity for CcpA. Next, the concentration of CcpA in the lane was gradually reduced to put the probes in a semi-binding state. When using 0.9 μM CcpA in the lane, a high affinity for CcpA was shown in T4, T5, T6, T9, and T10, indicating lower conservation for these five positions. In contrast, a low affinity for CcpA was shown in T1, T2, T3, T7, T8, T11, and T12, indicating that these seven positions are crucial for CcpA binding (Figures 3B and 3C). A further mutation in the symmetrical region showed that the affinity of probes for CcpA was weakened or disappeared, indicating that the symmetrical region in CRE_{Tre} was not allowed to be nonmatched in more positions (Figure S3).

Inversion of symmetric regions leads to the transition from CCA to CCR

Next, pBLTE was constructed with a reporter protein eGFP controlled by P*treA* to screen different carbon sources for significant effectors corresponding to either CCR or CCA. Glucose, fructose, mannitol, glycerol, sucrose, mannose, sorbitol, arabinose, xylose, and trehalose were added at concentrations of 1.5%. The results showed that glucose, fructose, mannitol, sucrose, mannose, and xylose had negative effects on eGFP production. Of these carbohydrates, glucose exhibited the greatest inhibition (Figure 4A). Glycerol, sorbitol, and arabinose had a promoting effect, and glycerol obtained the highest fluorescence intensity (Figure 4A). The formula I = (FI1-FI2)/FI1×100%, where FI1 represents the fluorescence intensity when only trehalose was added and FI2 represents the fluorescence intensity when trehalose and a specific carbohydrate were both added, was used to define the CCR/CCA effect. Therefore, $I_{glucose}$ was 84.35% and $I_{glycerol}$ was −35.72% (Figure 4B). As a result, glucose and glycerol were chosen as significant effectors for CCR and CCA, respectively.

Then, BlspT1E and BlspT2E were constructed with a reporter protein eGFP controlled by P*treA*-1 and P*treA*-2 to explore the relationship of the sequence of CRE with CCR/CCA. CRE_{Tre} (AGCTTT-AT-AAAGCT) was used to replace the original CRE site (TGAAAGCGCTATAA) in P*treA*, generating P*treA*-1. Meanwhile, the symmetric region in CRE_{Tre} was inversed to generate $CRE_{Tre(R)}$ (AAAGCT-AT-AGCTTT). This synthetic CRE was used to replace the original CRE site in P*treA*, generating P*treA*-2. Two plasmids, pBLT1E and

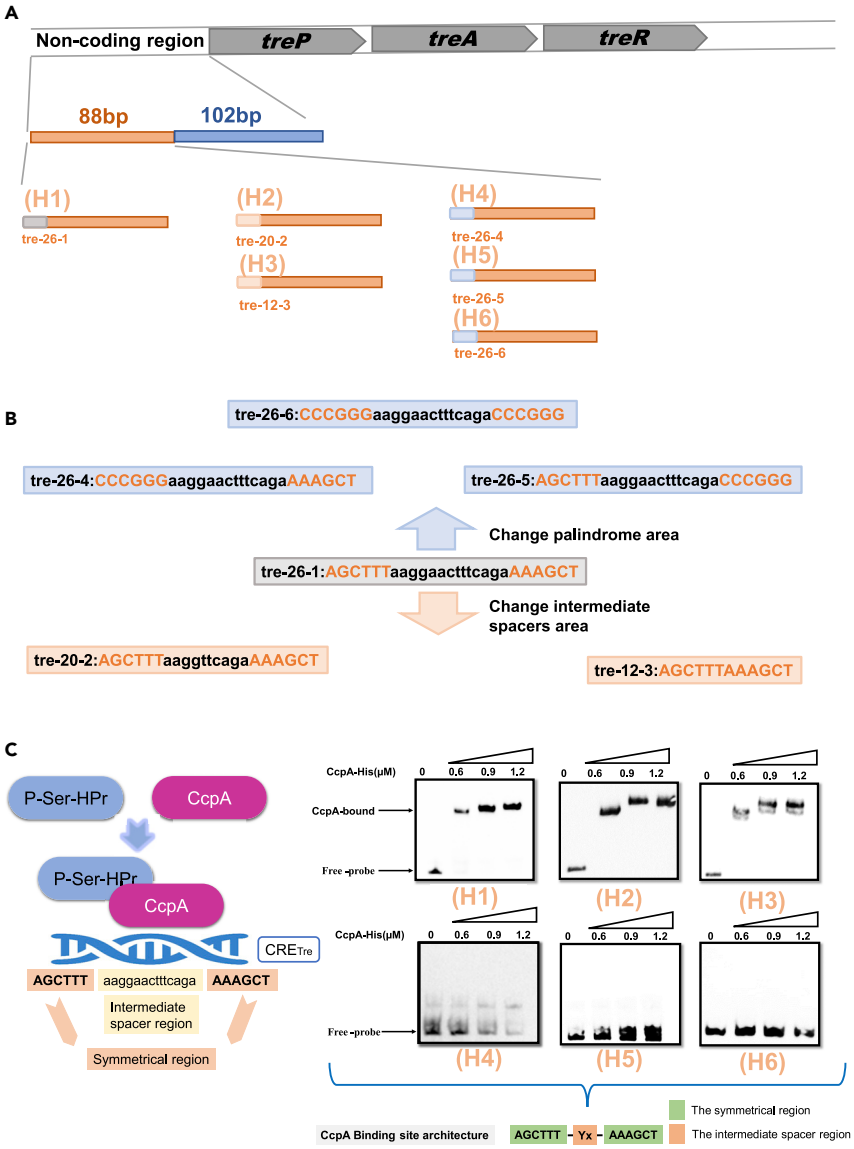


Figure 2. Influence of the 12-bp symmetrical region and the intermediate spacer region with CRE_{Tre} (AGCTTT-Yx-AAAGCT) on CcpA protein regulation

(A) Construction of recombination fragments (fragment H1, fragment H2, fragment H3, fragment H4, fragment H5, and fragment H6) harboring the CRE_{Tre} with different intermediate spacers length or different 12-bp symmetrical region.

(B) Two fragments that change the intermediate spacer region and three fragments that change the 12-bp symmetrical region are derived from tre-26-1 fragment, the intermediate spacer region with black and the 12-bp symmetrical region with red.

(C) EMSA of CcpA protein binding to six fragments (fragment H1, fragment H2, fragment H3, fragment H4, fragment H5, and fragment H6) that carrying different CRE_{Tre} sites.

pBLT2E, which were carrying P*treA*-1 and P*treA*-2, were transformed to yield BlspT1E and BlspT2E. The above two plasmids were also transformed into *Bacillus licheniformis* CA, pBLT1E, and pBLT2E. Six strains, BlspTE, BlspTE1, BlspT1E, BlspT1E1, BlspT2E, and BlspT2E1 (Table S1), were cultured under different induction conditions: (1) 1.5% trehalose, (2) 1.5% trehalose +1.5% glucose, and (3) 1.5% trehalose +1.5% glycerol. For strains BlspTE and BlspT2E, the fluorescence intensity decreased when glucose was present (Figures 4C and 4E). Compared to the set without glucose, the fluorescence intensities decreased by 85.41% and 87.91%, indicating that CRE and $CRE_{Tre(R)}$ corresponded to the CCR effect mediated by glucose. In contrast, for strain BlspT1E, fluorescence intensity increased by 36.36% compared to that of

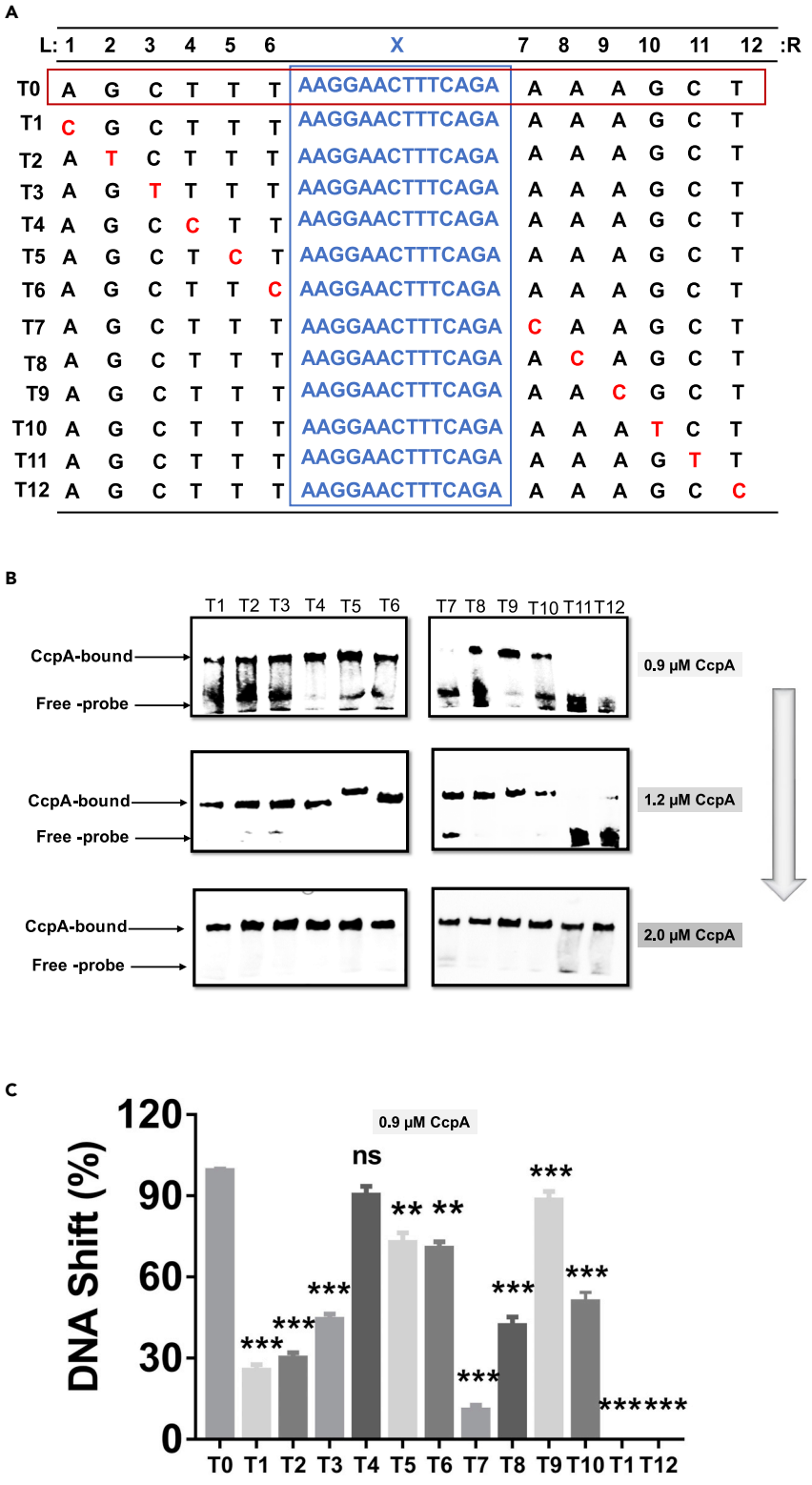


Figure 3. Influence of single point mutation of the 12-bp symmetrical region in CRE_{Tre} for CcpA protein binding

(A) Single point mutation of the 12-bp symmetrical region of CRE_{Tre}.

(B) EMSA of CcpA protein for 12 mutants (5′ biotin), 12 mutants (T1, T2, T3, T4, T5, T6, T7, T8, T9, T10, T11, and T12) derived from CRE_{Tre}, with concentrations of 0.9 μM–2.0 μM of CcpA protein.

Figure 3. *Continued*

(C) The ratio of protein-bound probe/total probe for 12 mutants that derived from CRE_{Tre} at the 0.9 μM CcpA protein was shown. Statistical significance was determined by Student's t-test (*P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001).

the set without glucose, indicating that CRE_{Tre} corresponded to the CCA effect in the presence of glucose (Figures 4D and 4F). We next investigated the effect of glycerol. For strains BlspTE and BlspT2E, the fluorescence intensity increased (Figures 4C and 4E). Compared to the set without glycerol, fluorescence intensities increased by 41.27% and 54.41%, respectively, indicating that CRE and $CRE_{Tre(R)}$ responded to the CCA effect caused by glycerol. In contrast, for strain BlspT1E, the fluorescence intensity decreased by 81.51% compared to the set without glycerol (Figures 4D and 4F), indicating that CRE_{Tre} corresponded to the CCR effect in the presence of glycerol. These results suggest that CRE and the synthetic CRE_{Tre} are opposite in regulatory direction (activation or inhibition). When the CcpA protein was deleted, the CCA effect or the CCR effect in the strains BlspTE1, BlspT1E1, and BlspT2E1 was weakened or even disappeared (Figures 4C–4E). This result also indicates that the CCA effect and the CCR effect in *B. licheniformis* are mainly mediated by the CcpA protein. Therefore, we can conclude that the fine-tuning

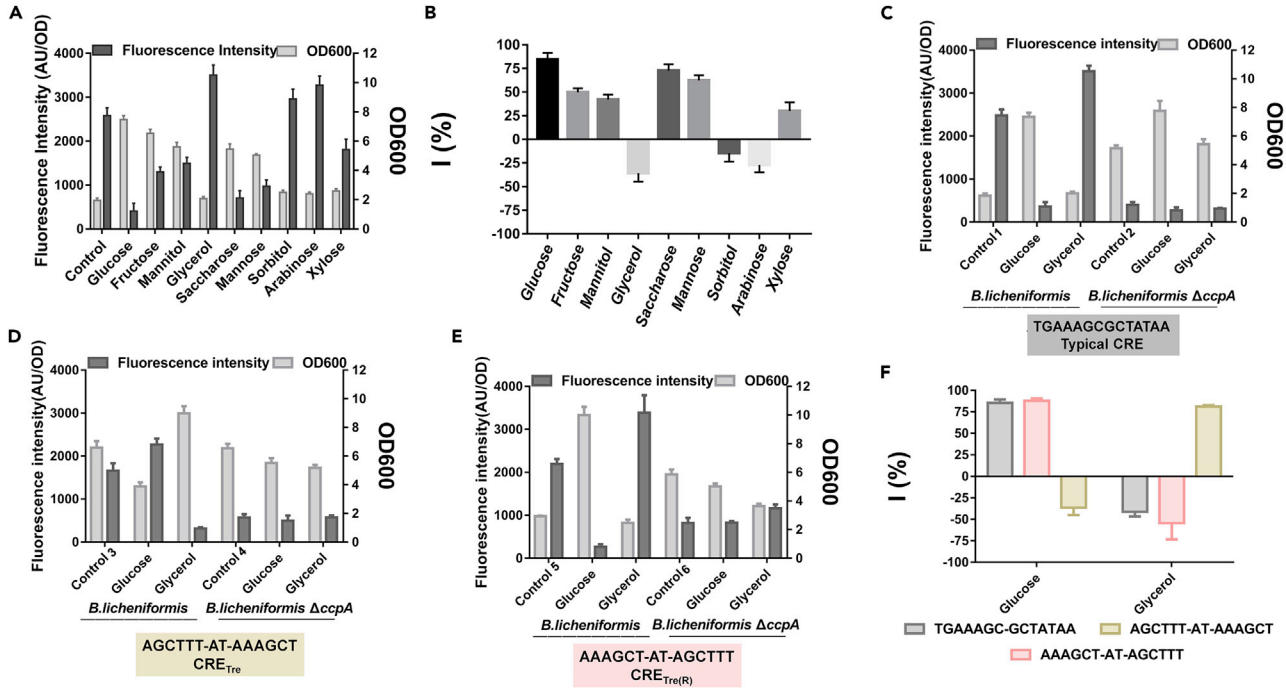


Figure 4. Characterization of the CcpA regulation *in vivo* for trehalose-inducible system

(A) The fluorescence intensity and OD 600 were measured for the strain carrying pBLTE plasmid under ten conditions (control: 1.5% trehalose, 1.5% trehalose + 1.5% glucose, 1.5% trehalose + 1.5% fructose, 1.5% trehalose + 1.5% mannitol, 1.5% trehalose + 1.5% glycerol, 1.5% trehalose + 1.5% glycerol, 1.5% trehalose + 1.5% saccharose, 1.5% trehalose + 1.5% mannose, 1.5% trehalose + 1.5% sorbitol, 1.5% trehalose + 1.5% arabinose, 1.5% trehalose + 1.5% xylose). Data are shown as means ± SD, n = 3.

(B) Quantify the CCR/CCA effect with the formula $I=(FI_1-FI_2)/FI_1\times100\%$ due to different carbohydrates (FI_1 represents the fluorescence intensity when only trehalose is added. FI_2 represents the fluorescence intensity when trehalose is added and another carbohydrate is also added. Glucose, fructose, mannitol, glycerol, sucrose, mannose, sorbitol, arabinose, and xylose added at a concentration of 1.5% while adding 1.5% trehalose. Data are shown as means ± SD, n = 3.

(C) The fluorescence intensity and OD 600 were measured for trehalose promoter in both the strain BlspTE (control 1: 1.5% trehalose, 1.5% trehalose + 1.5% glucose, 1.5% trehalose + 1.5% glycerol) and CcpA-deletion strain BlspTE1 (control 2: 1.5% trehalose, 1.5% trehalose + 1.5% glucose, 1.5% trehalose + 1.5% glycerol). Data are shown as means ± SD, n = 3.

(D) The fluorescence intensity and OD 600 were measured for trehalose promoter whose CRE site was replaced by CRE_{Tre} site in both the strain BlspT1E (control 3: 1.5% trehalose, 1.5% trehalose + 1.5% glucose, 1.5% trehalose + 1.5% glycerol) and CcpA-deletion strain BlspT1E1 (control 4: 1.5% trehalose, 1.5% trehalose + 1.5% glucose, 1.5% trehalose + 1.5% glycerol). Data are shown as means ± SD, n = 3.

(E) The fluorescence intensity and OD 600 were measured for trehalose promoter whose CRE site was replaced by $CRE_{Tre(R)}$ site in both the strain BlspT2E (control 5: 1.5% trehalose, 1.5% trehalose + 1.5% glucose, 1.5% trehalose + 1.5% glycerol) and CcpA-deletion strain BlspT2E1 (control 6: 1.5% trehalose, 1.5% trehalose + 1.5% glucose, 1.5% trehalose + 1.5% glycerol). Data are shown as means ± SD, n = 3.

(F) Compared three CRE sites by using formula $I=(FI_1-FI_2)/FI_1\times100\%$ while extra adding glucose or glycerol. Data are shown as means ± SD, n = 3.

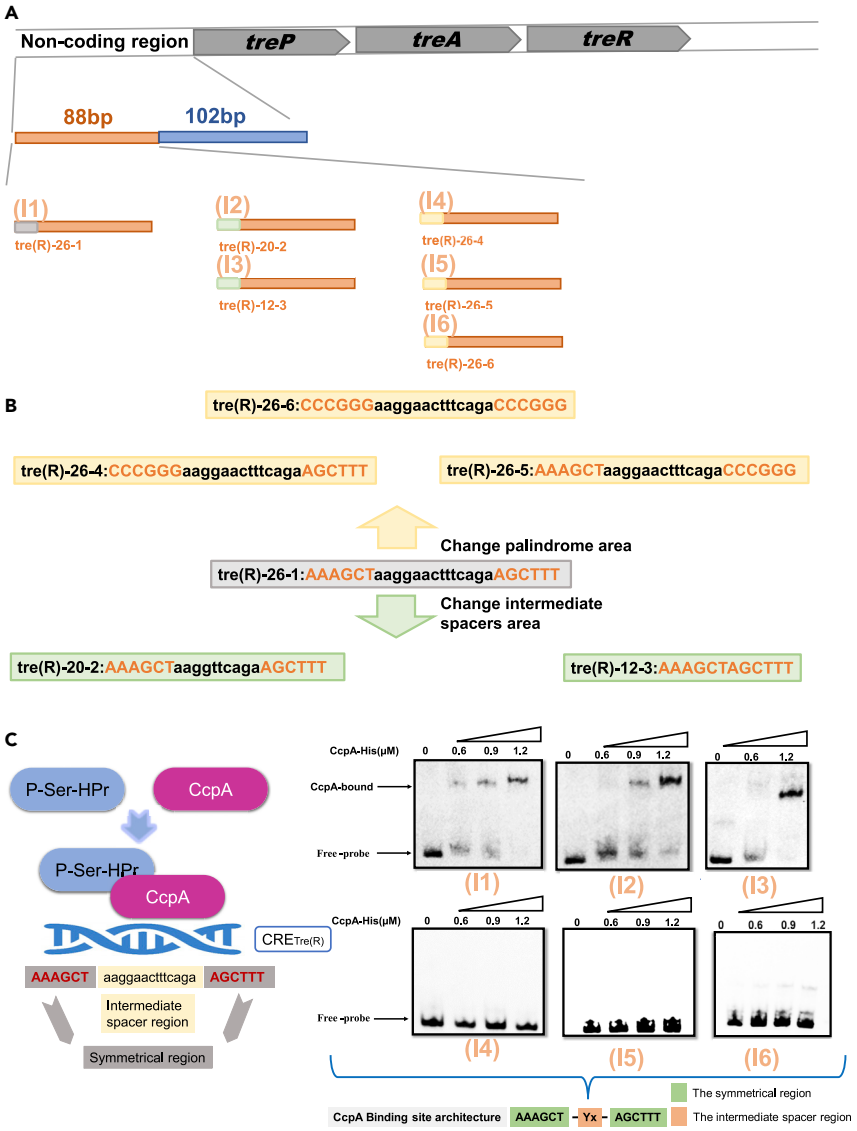


Figure 5. Influence of the 12-bp symmetrical region and the intermediate spacer region with $CRE_{Tre(R)}$ (AAAGCT-Yx-AGCTTT) on CcpA protein regulation

(A) Construction of recombination fragments (fragment I1, fragment I2, fragment I3, fragment I4, fragment I5, and fragment I6) harboring the $CRE_{Tre(R)}$ with different intermediate spacer length or different 12-bp symmetrical region.

(B) Two fragments that change the intermediate spacer region and three fragments that change the 12-bp symmetrical region are derived from tre(R)-26-1 fragment, the intermediate spacer region with black and the 12-bp symmetrical region with red.

(C) EMSA of CcpA protein binding to six fragments (fragment I1, fragment I2, fragment I3, fragment I4, fragment I5, and fragment I6) that carry different $CRE_{Tre(R)}$ sites.

of CcpA-mediated regulation may be modulated by the sequence of the CRE and the availability of the extracellular carbon source.

Characterization of the CcpA binding motif $CRE_{Tre(R)}$

The above results show that CRE_{Tre} and $CRE_{Tre(R)}$ exhibit opposite directions in CcpA-mediated regulatory *in vivo*. It is necessary to explore whether this difference is due to a varied affinity of CcpA. First, the role of the intermediate spacer region was confirmed. Three fragments, I1, I2, and I3, were constructed by changing the length of the intermediate spacer, as shown in Figures 5A and 5B. The EMSA results show that CcpA

can still bind to all of them (Figure 5C), even to the one with the intermediate spacer completely removed. This result suggests that CcpA-$CRE_{Tre(R)}$ interaction is independent of the intermediate spacer. Next, the role of 6-bp palindromes was investigated. The palindromic sequence in the 5′ end "AAAGCT" was replaced by "CCCGGG" to yield I4. In the 3′ end, "AGCTTT" was replaced by "CCCGGG" to yield I5. For I6, both palindromic sequences were replaced by "CCCGGG" and "CCCGGG", respectively (Figures 5A and 5B). The EMSA results showed that CcpA could no longer bind to I4, I5, or I6 (Figure 5C), indicating that these palindromic regions were crucial for CcpA binding. Thus, it can be hypothesized that the motif of $CRE_{Tre(R)}$ is "AAAGCT-Yx-AGCTTT" (Y stands for any base, and X stands for the number of bases).

Then, the different roles of the specific bases within the palindromic regions were characterized. For this purpose, 12 mutants (TR1, TR2, TR3, TR4, TR5, TR6, TR7, TR8, TR9, TR10, TR11, and TR12) (Figure 6A) were constructed based on fragment I1. Under a CcpA protein concentration of 1.2 μM, TR4, TR6, TR8, TR9, TR10, and TR12 still exhibited high affinity for CcpA, suggesting that these positions could tolerate substantial degeneracy. On the other hand, binding affinity was strongly reduced for mutants TR1, TR2, TR3, TR5, TR7, and TR11. The high conservatism of these positions suggests their important functional role (Figures 6B and 6C). Further mutations within the symmetric region show that the affinity of these probes for CcpA protein could be also diminished or even eliminated (Figure S4). These results indicate that the symmetric region could not tolerate more non-match bases.

Wide occurrence of the novel binding site CRE_{Tre} and $CRE_{Tre(R)}$ in *B. licheniformis*

To further confirm the newly found motif, two genes containing CRE_{Tre} or $CRE_{Tre(R)}$-like sequence were cloned. The former (AGCTTT-tggatttg-AAAGCT) was located in the 5′ non-coding regions of the gene *dnaP* (GenBank: NM_00300) (J) and the latter (AAAGCT-aaatgg-AGCTTT) was in gene 5A2 (GenBank: NM_09050) (K). As the EMSA results showed that CcpA protein can bind to J or K, the CRE-like sequence in both fragments was deleted, yielding J-Δcre and K-Δcre. As expected, CcpA could no longer bind to J-Δcre or K-Δcre (Figure 7B), suggesting the existence of CRE_{Tre} and $CRE_{Tre(R)}$. The strains pBLDE and pBLAE, which carried J and K as transcription initiation elements, were also constructed (Figure 7A). Both ensured the functional expression of eGFP. However, when the strains were subjected to cultivation under glucose stress (30 g/L), the fluorescence intensity of BlspDE increased by 273.98% and that of BlspAE decreased by 22.64%. These results further proved that CRE_{Tre} is a CCA-responsive element under glucose stress, while CRE_{Tre} is a CCR-responsive element (Figure 7C). Furthermore, the *B.licheniformis* genome was selected for CRE_{Tre} and $CRE_{Tre(R)}$, and 279 possible sequences were found distributed in 264 genes (Figure 8, Table S3). From the function of genes, these genes are mainly related to the following aspects: (1) carbohydrate metabolism, (2) amino acid metabolism, (3) sporulation formation, (4) bacterial migration, (5) transcription factors, (6) transporters, (7) DNA repair, recombination, (8) proteases, peptidases, hydrolases, (9) redox balance, and (10) unknown.

Four degenerate CRE_{Tre} sites were found in the xylose operon and mannose operon by alignment. The further EMSA results proved that CcpA can bind to $xylB^{*410}$ and $xylB^{*1384}$, while the other two, $xylR^{*775}$ and $xylA^{*249}$, had no affinity for CcpA. $xylR^{*775}$ and $xylA^{*249}$ both had a signal base mutation in the 5′ end ("T"→"C"), which could explain why these CRE_{Tre}-like sequences had no interaction with CcpA (Figure S5). Two degenerate CRE_{Tre} sites were found in the mannose operon, and the EMSA result showed that the CcpA protein can bind to both of the fragments (Figure S5).

DISCUSSION

CcpA participates in the regulation of central carbon and nitrogen metabolism (Yoshida et al., 2001; Halsey et al., 2017), the formation of biofilms (Römling and Galperin, 2015), the formation of spores (Voltersen et al., 2018), physiological processes (Charbonnier et al., 2017), DNA replication (Jolly et al., 2020), and the CCR effect. The interaction between this protein and target genes is the key molecular basis in the above process. In the current study, a new CcpA protein binding site in the *treR* gene, "AGCTTT-Yx-AAAGCT", called CRE_{Tre}, was identified. It is composed of two palindromic sequences and variable spacers. It has been proven that the DNA-binding domain contains a helix-turn-helix motif that makes base-specific contacts in the major groove of the DNA (Schumacher et al., 2004). In previous research, structure elucidation revealed that the CcpA-CRE complex varies in bending angle according to different CRE sequences (Schumacher et al., 1994). Compared to classical CRE sites with only highly conserved structures, the flexible structure of CRE_{Tre} provides new functionalities for CcpA-mediated regulation. A similar motif for protein recognition sites has been reported in *E. coli*. The transcription

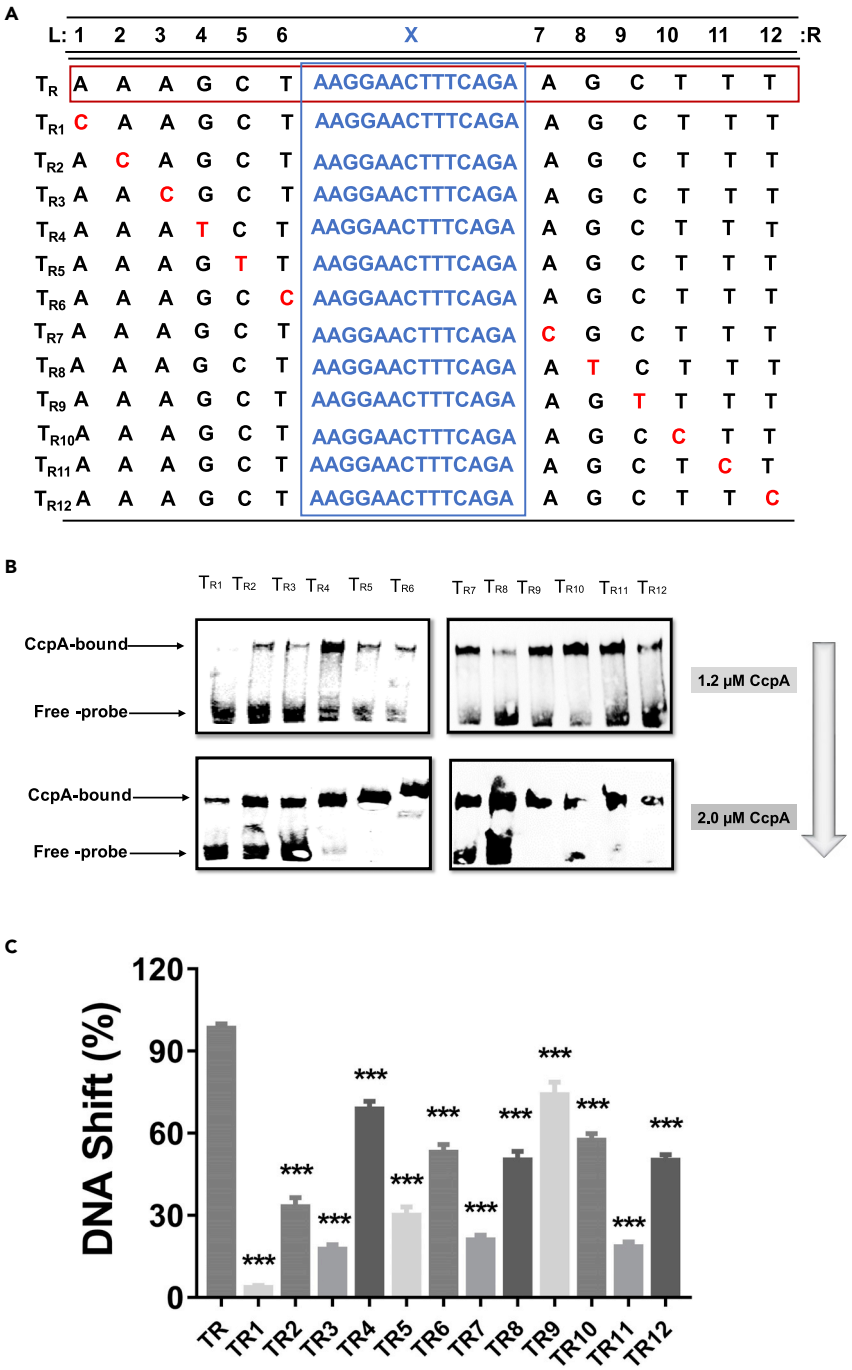


Figure 6. Influence of single point mutation of the 12-bp symmetrical region in $CRE_{Tre(R)}$ for CcpA protein binding

(A) Single point mutation of the 12-bp symmetrical region of $CRE_{Tre(R)}$.

(B) EMSA of CcpA protein for 12 mutants (5′ biotin), 12 mutants (TR1, TR2, TR3, TR4, TR5, TR6, TR7, TR8, TR9, TR10, TR11, and TR12) derived from $CRE_{Tre(R)}$, with concentrations of 1.2 μM–2.0 μM of CcpA protein.

(C) The ratio of protein-bound probe/total probe for 12 mutants that derived from $CRE_{Tre(R)}$ at the 1.2 μM CcpA protein were shown. Statistical significance was determined by Student's t-test (*P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001).

factor HipB can recognize a motif consisting of a spacer sequence and a palindromic sequence. The crystal structure revealed that the protein can bind to the motif by extruding the spacer (Schumacher et al., 2015; Lin et al., 2013a). However, the functional properties of the spacer sequence remained unknown.

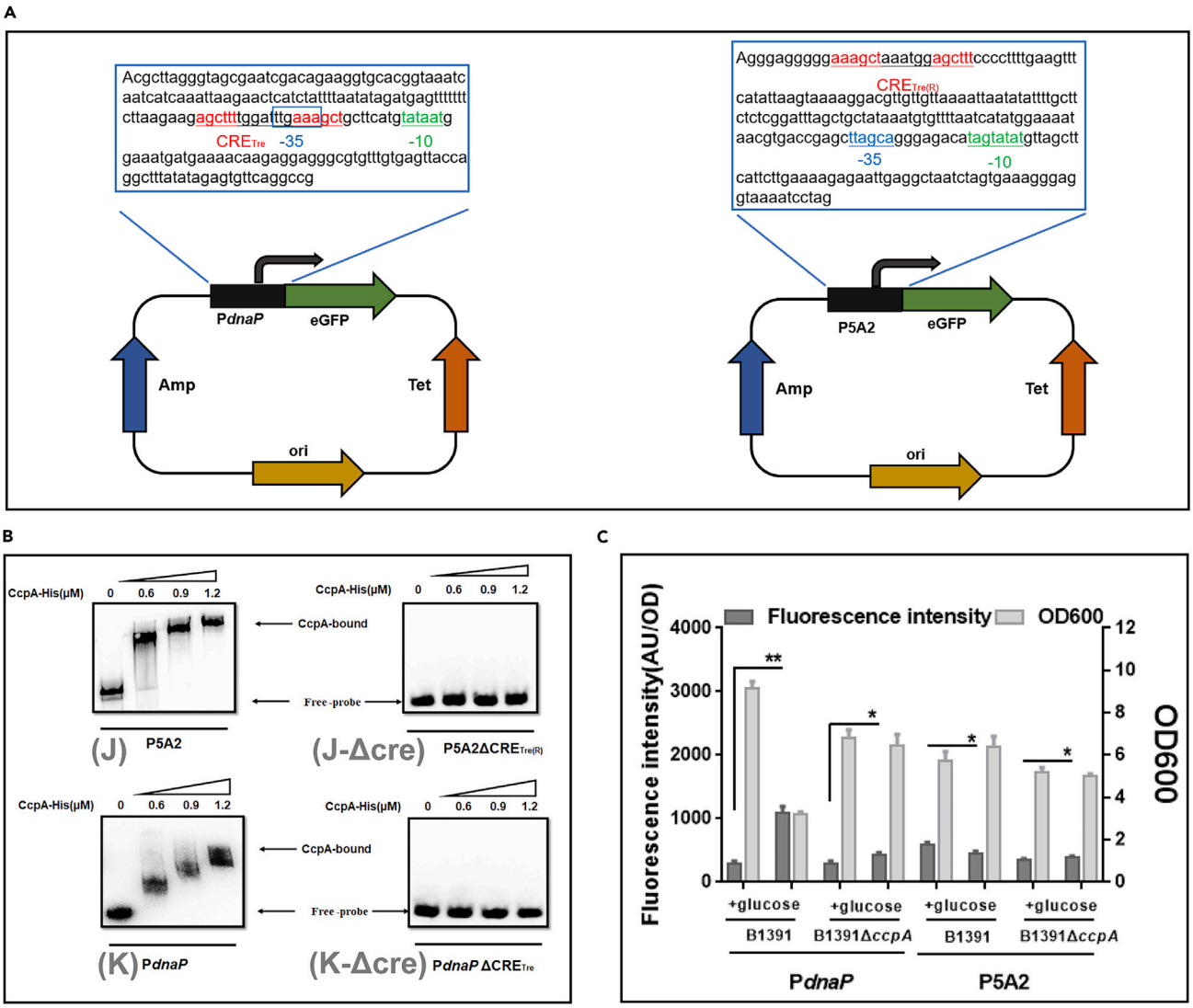


Figure 7. Exploring the property of CRE_{Tre} in P*danp*, the $CRE_{Tre(R)}$ in P5A2

(A) The sequences of the P*dnaP* and P5A2 from *Bacillus licheniformis* are shown, and the CRE_{Tre} or $CRE_{Tre(R)}$ is shown in red, −10 region is shown in green, and −35 region is shown in blue.

(B) EMSA of CcpA protein binding to two fragments (P*dnaP*, P5A2) labeled with 5′ biotin, while not binding to two fragments (P*dnaP*ΔCRE_{Tre}, P5A2Δ$CRE_{Tre(R)}$) labeled with 5′ biotin.

(C) The OD 600 and the fluorescence intensity were measured in both the *Bacillus licheniformis* CICIM B1391 and CcpA-deletion strain when using P*dnaP* or P5A2 as the expression element. Data are shown as means ± SD, n = 3. Statistical significance was determined by Student's t-test (*P ≤ 0.05; **P ≤ 0.01; ***P ≤ 0.001).

This study confirms that although the palindromic sequence was the indispensable element of the recognition site, modification in the variable spacers was able to modulate the interaction between CRE_{Tre} and CcpA. This result provides new insight into the mechanism by which the regulator interacts with the target genes, suggesting that the spacer sequence within this motif may serve as a target for the fine-tuning of the regulation.

The CcpA protein can bind to different target sites, resulting in either a CCR effect or a CCA effect. However, the relationship between the target sequence feature and the direction of CcpA-mediated regulation is poorly understood. This study reported an interesting new motif, CRE_{Tre}, which is involved in a bidirectional regulation in the glucose stress environment. Previous studies mainly focused on substrates corresponding to CCR or CCA. For example, glucose-related CCR effects have been investigated extensively. The CCR

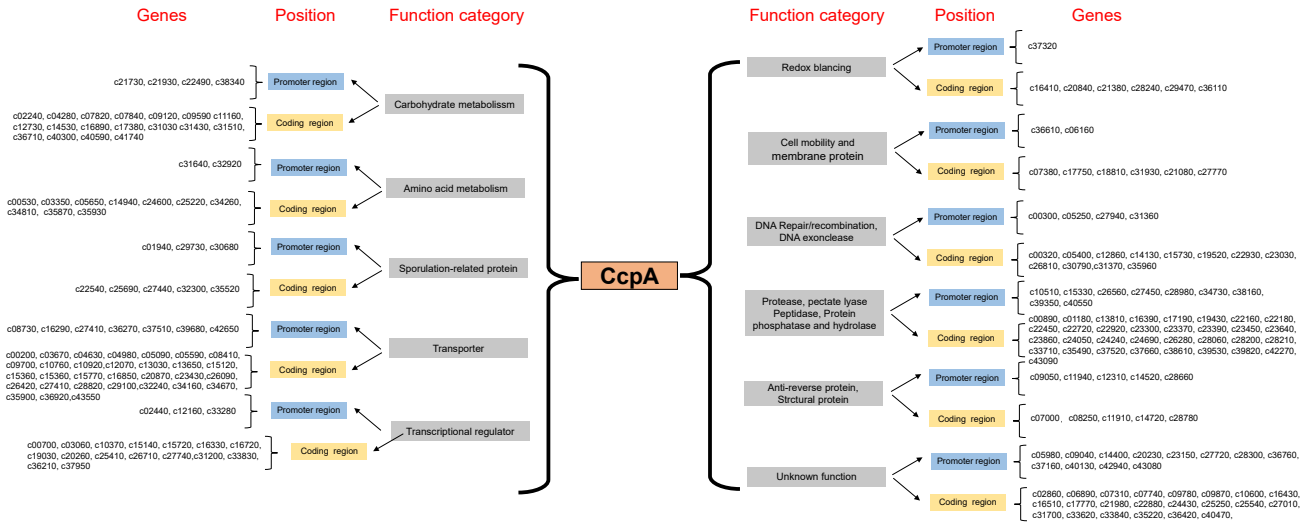


Figure 8. Searching *Bacillus licheniformis* genome for CRE_{Tre} and $CRE_{Tre(R)}$

effect or CCA effect is also caused by the different priority utilization levels of the carbon source. When a rapid-acting carbon source (glucose) is present in a fermentation system, the utilization priority of the rapid-acting carbon source is higher than that of the inducer, causing the inducer (another sugar or alcohol)-related genes to be inhibited by the CcpA protein and subjecting the expression system to the CCR effect (Jankovic and Brückner, 2007; Moreno et al., 2010). When a delayed carbon source is present in a fermentation system, the inducer (sugar or alcohol) is the fast-acting carbon source, and the related metabolic utilization genes involved are activated by the CcpA protein, thereby causing the expression system to be subjected to the CCA effect (Yu et al., 2017). Sugar/alcohol-inducible promoters usually contain a classical CRE site in the core region (Heravi et al., 2011; Ren et al., 2010), which makes the regulation of the CcpA protein particularly important for the promoter. It is worth mentioning that the mutation of CcpA causes different effects for target genes (Hueck et al., 2010). Our study indicated that the transition between CCR and CCA is made possible by rearrangement of the nucleotide sequence of CRE_{Tre}. Because of its indispensable role in the CcpA-mediated CCR effect, the classical CRE can be likened to a molecular switch that turns the regulation on and off. The newly found CRE_{Tre} exhibits competence in this two-way molecular switch, guiding a regulation with either CCR or CCA effect. We also used both constitutive and inducible promoters as examples to further explore the influence of CRE_{Tre} or $CRE_{Tre(R)}$ on transcription. The results were consistent with those in the trehalose operon. This CRE-engineering strategy has great potential in synthetic biology, particularly for genetic element development.

From the distribution of CRE_{Tre} and $CRE_{Tre(R)}$ in *B. licheniformis*, it was found that the genes present in CRE_{Tre} and $CRE_{Tre(R)}$ are relevant to important biological activities in bacteria, such as DNA mismatch and repair genes, sporulation genes, cell movement and division genes, sugar alcohol transporters, and ion transporters, which are responsible for the PTS system. It is worth mentioning that CRE_{Tre} and $CRE_{Tre(R)}$ are also present in toxin-encoding genes and promoters, such as subtilisin (GenBank: NM_12310) and autolysin inhibitor (GenBank: NM_14520). This fact indicates that the CcpA protein still has some unknown functions, and it shows the importance of CRE_{Tre} and $CRE_{Tre(R)}$ as cis-acting elements. CRE_{Tre} and $CRE_{Tre(R)}$ have opposite functional properties, and there is also some symmetry in their sequences. The result of a single base mutation on the two 6-bp palindromic sequences of CRE_{Tre} and $CRE_{Tre(R)}$ indicates that a single base mutation is allowed, but further mutations after a single base mutation result in the loss of CcpA protein binding ability, which indicates that CRE_{Tre} and $CRE_{Tre(R)}$ allow some changes for two 6-bp palindromic sequences, but this change will affect the binding of the CcpA protein to the site. CRE_{Tre} and $CRE_{Tre(R)}$ provide two effective pathways for the regulation of the CcpA protein in both function and sequence.

In conclusion, in addition to more research on carbon metabolism via the CcpA protein, further research is needed on other aspects of its function. The discovery and function of the CRE_{Tre} and $CRE_{Tre(R)}$ sites further

clarify the regulatory network of CcpA proteins. It is conceivable that the functions and sites play important roles in the CcpA protein-mediated regulatory network.

Limitation of the study

Here, we show the CCA/CCR effect with the newly CcpA binding site under different carbon sources. Notably, there also exists unknown CRE site in *B.licheniformis*. Further studies would mine the potential CRE sites in *B.licheniformis*.

Resource availability

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Prof. Guiyang Shi, gyshi@jiangnan.edu.cn.

Material availability

All tables and figures are included in the text and supplemental information.

Data and code availability

The published article contains all data generated or analyzed.

METHODS

All methods can be found in the accompanying Transparent Methods supplemental file.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2021.102400.

ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (2020YFA0907700, 2018YFA0900300, 2018YFA0900504 and 2016YFD0401404), the National Natural Foundation of China (31401674), the National First-Class Discipline Program of Light Industry Technology and Engineering (LITE2018-22), and the Top-notch Academic Programs Project of Jiangsu Higher Education Institutions.

AUTHOR CONTRIBUTIONS

F.X., Y.L., and G.S. designed the study, carried out the experiments, analyzed data, and wrote the paper. Y.Z. and H.W. carried out the experiments. L.Z. and Z.D. analyzed data. Z.G. and S.X. designed the study.

DECLARATION OF INTERESTS

The authors declare that they have no conflicts of interest.

Received: December 17, 2020
Revised: March 6, 2021
Accepted: April 5, 2021
Published: May 21, 2021

REFERENCES

Antunes, A., Camiade, E., Monot, M., Courtois, E., Barbut, F., Sernova, N.V., Rodionov, D.A., Martin-Verstraete, I., and Dupuy, B. (2012). Global transcriptional control by glucose and carbon regulator CcpA in Clostridium difficile. Nucleic Acids Res. *40*, 10701–10718.

Asai, K., Baik, S.H., Kasahara, Y., Moriya, S., and Ogasawara, N. (2000). Regulation of the transport system for C4-dicarboxylic acids in *Bacillus subtilis*. Microbiology *146*, 263–271.

Bruckner, R., and Titgemeyer, F. (2002). Carbon catabolite repression in bacteria: choice of the carbon source and autoregulatory limitation of sugar utilization. FEMS Microbiol. Lett. *209*, 141–148.

Charbonnier, T., LeCoq, D., McGovern, S., Calabre, M., Delumeau, O., Aymerich, S., and Jules, M. (2017). Molecular and physiological logics of the pyruvate-induced response of a novel transporter in *Bacillus subtilis*. mBio *8*, e0097617.

Eastmond, P.J., and Graham, I.A. (2003). Trehalose metabolism: a regulatory role for trehalose-6-phosphate? Curr. Opin. Plant Biol. *6*, 231–235.

Fujita, Y. (2014). Carbon catabolite control of the metabolic network in *Bacillus subtilis*. Biosci. Biotechnol. Biochem. *3*, 245–259.

Görke, B., and Stülke, J. (2008). Carbon catabolite repression in bacteria: many ways to make the

most out of nutrients. Nat. Rev. Microbiol. *6*, 613–624.

Gosseringer, R., Küster, E., Galinier, A., Deutscher, J., and Hillen, W. (1997). Cooperative and non-cooperative DNA binding modes of catabolite control protein CcpA from *Bacillus megaterium* result from sensing two different signals. J. Mol. Biol. *266*, 665–676.

Halsey, C.R., Lei, S., Wax, J.K., Lehman, M.K., Nuxoll, A.S., Steinke, L., Sadykov, M., Powers, R., and Fey, P.D. (2017). Amino acid catabolism in *Staphylococcus aureus* and the function of carbon catabolite repression. mBio *8*, e0143416.

Heravi, K.M., Wenzel, M., and Altenbuchner, J. (2011). Regulation of *mtl* operon promoter of *Bacillus subtilis*: requirements of its use in expression vectors. Microb. Cell Fact. *10*, 83.

Hueck, C.J., Kraus, A., Schmiedel, D., and Hillen, W. (2010). Cloning, expression and functional analyses of the catabolite control protein CcpA from *Bacillus megaterium*. Mol. Microbiol. *16*, 855–864.

Jankovic, I., and Brückner, R. (2007). Carbon catabolite repression of sucrose utilization in *Staphylococcus xylosus*: catabolite control protein CcpA ensures glucose preference and autoregulatory limitation of sucrose utilization. J. Mol. Microb. Biotech. *12*, 114–120.

Jolly, S.M., Gainetdino, I., Jouravleva, K., Zhang, H., Strittmatter, L., Bailey, S.M., Hendricks, G.M., Dhabaria, A., Ueberheide, B., and Zamore, P.D. (2020). *Thermus thermophilus* argonaute functions in the completion of DNA replication. Cell *182*, 1545–1559.

Jörg, S., and Hillen, W. (1999). Carbon catabolite repression in bacteria. Curr. Opin. Microbiol. *2*, 195–201.

Kim, J.N., and Burne, R.A. (2017). CcpA and CodY coordinate acetate metabolism in *Streptococcus mutans*. Appl. Environ. Microbiol. *83*, e03274–16.

Kleijn, R.J., Buescher, J.M., Chat, L.L., Jules, M., Aymerich, S., and Saue, U. (2010). Metabolic fluxes during strong carbon catabolite repression by malate in Bacillus subtilis. J. Biol. Chem. *285*, 1587–1596.

Li, Y.R., Jin, K., Zhang, L., Ding, Z.Y., Gu, Z.H., and Shi, G.Y. (2018). Development of an inducible secretory expression system in *Bacillus licheniformis* based on an engineered xylose operon. J. Agric. Food Chem. *66*, 9456–9464.

Li, Y.R., Liu, X., Zhang, L., Ding, Z., and Shi, G.Y. (2019). Transcriptional changes in the xylose operon in *Bacillus licheniformis* and their use in fermentation optimization. Int. J. Mol. Sci. *20*, 4615.

Lin, C.Y., Awano, N., Masuda, H., Park, J.H., and Inouye, M. (2013a). Transcriptional repressor HipB regulates the multiple promoters in *Escherichia coli*. J. Mol. Microbiol. Biotechnol. *23*, 440–447.

Lin, Z., Chul, C.S., Danko, C.G., Adam, S., Stanhope, M.J., and Burne, R.A. (2013b). Gene regulation by CcpA and catabolite repression explored by RNA-seq in *Streptococcus mutans*. PLoS One *8*, e60465.

Loll, B., Alings, C., Kowalczyk, M., Chieduch, A., Bardowski, J., Saenger, W., and Biesiadka, J. (2007). Structure of the transcription regulator CcpA from *Lactococcus lactis*. Acta Crystallogr. D *63*, 431–436.

Moreno, M.S., Schneider, B.L., Maile, R.R., Weyler, W., and Saier, M.H. (2010). Catabolite repression mediated by the CcpA protein in *Bacillus subtilis*: novel modes of regulation revealed by whole-genome analyses. Mol. Microbiol. *39*, 1366–1381.

Okabe, S., Shafdar, A.A., Kobayashi, K., Zhang, L., and oshiki, M. (2020). Glycogen metabolism of the anammox bacterium "*Candidatus* Brocadia sinica". ISME J. https://doi.org/10.1038/s41396-020-00850-5.

Patricia, F.R., Ipsita, R., Mark, O., and Tajalli, K. (2014). Proteomics analysis of *Bacillus licheniformis* in response to oligosaccharides elicitors. Enzyme Microb. Tech. *61*, 61–66.

Ren, C., Gu, Y., Hu, S., Wu, Y., Wang, P., Yang, Y.L., Yang, C., Yang, S., and Jiang, W.H. (2010). Identification and inactivation of pleiotropic regulator CcpA to eliminate glucose repression of xylose utilization in *Clostridium acetobutylicum*. Metab. Eng. *12*, 446–454.

Römling, U., and Galperin, M.Y. (2015). Bacterial cellulose biosynthesis: diversity of operons, subunits, products, and functions. Trends Microbiol. *23*, 545–557.

Ruud, D.O.W., Gerald, S., and Kuipers, O.P. (2015). Probing the regulatory effects of specific mutations in three major binding domains of the pleiotropic regulator CcpA of *Bacillus subtilis*. Front. Microbiol. *6*, 1051.

Schöck, F., and Dahl, M.K. (1996). Expression of the tre operon of *Bcillus subtilis*168 is regulated by the repressor TreR. J. Biotechnol. *178*, 4576–4581.

Schumacher, M.A., Choi, K.Y., Zalkin, H., and Brennan, R.G. (1994). Crystal structure of LacI member PurR bound to DNA: minor groove binding by alpha helices. Science *266*, 763–770.

Schumacher, M.A., Allen, G.S., Diel, M., Seidel, G., Hillen, W., and Brennan, R.G. (2004). Structural basis for allosteric control of the transcription regulator CcpA by the phosphoprotein HPr-Ser46-P. Cell *118*, 731–741.

Schumacher, M.A., Sprehe, M., Bartholomae, M., Hillen, W., and Brennan, R.G. (2011). Structures of carbon catabolite protein A-(Hpr-Ser46-P) bound to diverse catabolite response element sites reveal the basis for high-affinity binding to degenerate DNA operators. Nucleic Acids Res. *39*, 2931–2942.

Schumacher, M.A., Balani, P., Min, J.K., Chinnam, N.B., Hansen, S., Vulic, M., Lewis, K., and Brennan, R.G. (2015). HipBA-promoter structures reveal the basis of heritable multidrug tolerance. Nature *524*, 59–64.

Shi, J., Zhan, Y.Y., Zhou, M.L., He, M., Wang, Q., Li, X., Wen, Z.Y., and Chen, S.W. (2019). High-level production of short branched-chain fatty acids from waste materials by genetically modified *Bacillus licheniformis*. Bioresour. Technol. *271*, 325–331.

Sonenshein, A. (2007). Control of key metabolic intersections in *Bacillus subtilis*. Nat. Rev. Microbiol. *5*, 917–927.

Voltersen, V., Blango, M.G., Herrmann, S., Schmidt, F., Heinekamp, T., Strassburger, M., Krüger, T., Bacher, P., Lother, J., Weiss, E., et al. (2018). Proteome analysis reveals the conidial surface protein CcpA essential for virulence of the pathogenic fungus *Aspergillus fumigatus*. mBio *9*, e01557–18.

Warner, J.B., and Lolkema, J.S. (2003). CcpA-dependent carbon catabolite repression in bacteria. Microbiol. Mol. Biol. Rev. *67*, 475–490.

Weickert, M.J., and Chambliss, G.H. (1990). Site-directed mutagenesis of a catabolite repression operator sequence in *Bacillus subtilis*. Proc. Natl. Acad. Sci. U S A *87*, 6238–6242.

Willenborg, J., DeGreeff, A., Jarek, M., Valentin-Weigand, P., and Goethe, R. (2014). The CcpA regulon of *Streptococcus suis* reveals novel insights into the regulation of the streptococcal central carbon metabolism by binding of CcpA to two distinct binding motifs. Mol. Microbiol. *92*, 61–83.

Xiao, Z., and Xu, P. (2007). Acetoin metabolism in bacteria. Crit. Rev. Microbiol. *33*, 127–140.

Xiao, F.X., Li, Y.R., Zhang, Y.P., Wang, H.,R., Zhang, L., Ding, Z.Y., Gu, Z.H., Xu, S., and Shi, G.Y. (2020). Construction of a novel sugar alcohol-inducible expression system in *Bacillus licheniformis*. Appl. Microbiol. Biotechnol. *104*, 5409–5425.

Yoshida, K.I., Kobayashi, K., Miwa, Y., Kang, C.M., Matsunaga, M., Yamaguchi, H., Tojo, S., Yamamoto, M., Nishi, R., Ogasawara, N., et al. (2001). Combined transcriptome and proteome analysis as a powerful approach to study genes under glucose repression in Bacillus subtilis. Nucleic Acids Res. *29*, 683–692.

Yu, W., Chen, Z., Ye, H., Liu, P., Li, Z., and Wang, Y. (2017). Effect of glucose on poly-γ-glutamic acid metabolism in *Bacillus licheniformis*. Microb. Cell Fact. *16*, 22.

Zhang, L., Liu, Y.Q., Yang, Y.P., Jiang, W.H., and Gu, Y. (2018). A novel dual-cre motif enables two-way autoregulation of CcpA in *Clostridium acetobutylicum*. Appl. Environ. Microbiol. *84*, e00114–e00118.

iScience, Volume *24*

Supplemental information

A new CcpA binding site plays a bidirectional role in carbon catabolism in *Bacillus licheniformis*

Fengxu Xiao, Youran Li, Yupeng Zhang, Hanrong Wang, Liang Zhang, Zhongyang Ding, Zhenghua Gu, Sha Xu, and Guiyang Shi

Supplementary Information

A newly CcpA binding site plays a bidirectional role in carbon catabolism in *Bacillus licheniformis*

Fengxu Xiao[a,b,c], Youran Li[a,b,c*], Yupeng Zhang[a,b,c], Hanrong Wang[a,b,c], Liang Zhang[a,b,c], Zhongyang Ding[a,b,c], Zhenghua Gu[a,b,c], Sha Xu[a,b,c], Guiyang Shi[a,b,c,d*]

[a]Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi 214122, People's Republic of China;

[b]National Engineering Laboratory for Cereal Fermentation Technology, Jiangnan University, 1800 Lihu Avenue, Wuxi, Jiangsu Province 214122, People's Republic of China;

[c]Jiangsu Provincial Research Center for Bioactive Product Processing Technology, Jiangnan University.

[d]Lead contact

*Address Correspondence to:

Youran Li, Jiangnan University, Wuxi, China. Email: liyouran@jiangnan.edu.cn

or:

Guiyang Shi, Jiangnan University, Wuxi, China. Email: gyshi@jiangnan.edu.cn

Legends:

Transparent Methods

Bacterial Strains, Plasmids, and Culture Conditions

Table S1 lists the strains and plasmids used or produced in this study. *Escherichia coli* JM109 was used for plasmid preparation. The *E*. *coli* JM109 transformants carrying plasmids based on pHY-PLK300, were selected on Luria–Bertani (LB) agar plates supplemented with ampicillin (100 μg/mL). The *B*. *licheniformis* B1391 transformants were selected on LB agar plates supplemented with tetracycline (20 μg/mL). Two strains were cultured in LB medium (10 g/L tryptone, 5 g/L yeast extract, 10 g/L NaCl), *E. coli* was cultured at 37 °C and 200 rpm, and *B. licheniformis* was cultured at 37 °C and 250 rpm. Fermentation medium (12 g/L tryptone, 24 g/L yeast extract, 16.427 g/L $K_2HPO_4.3H_2O$, 2.31 g/L KH_2PO_4) was prepared for protein expression. A 3% seed liquid was inoculated into the fermentation medium, and batch fermentation was cultured at 37 °C and 250 rpm.

Plasmid Construction

The recombinant plasmids were constructed based on pHY300-PLK using the primers listed in Table S2. First, eGFP was amplified by the primer pair eGFP-F/eGFP-R (Table S2). The fragment was then purified and digested with *Xho*I and *Sma*I, followed by incorporation into pHY300-PLK, yielding pE. Then, the promoter region, P*treA*, was cloned using 100 ng of the genome of *Bacillus licheniformis* CICIM B1391. PCR was performed with high-fidelity DNA polymerase (Vazyme Biotech Co., Ltd. 2 × Phanta Master Mix) and the primer pairs P*treA*-F/P*treA*-R. The fragment was also purified and digested with *Hin*dIII and *Xho*I, followed by incorporation into pE, yielding pBLTE. Next, the artificial promoter P*treA* was constructed using overlap extension PCR. Two fragments, P*treA*-CRETre-1 and P*treA*-CRETre-2, were cloned by the primer pairs P*treA*-F/ P*treA*-CRETre-R, P*treA*-CRETre-F/ P*treA*-R, using P*treA* as template. P*treA*-CRETre was created using P*treA*-CRETre-1 and P*treA*-CRETre-2 as a template, and P*treA*-F/P*treA*-R as primers. The fragment P*treA*-CRETre was also purified and digested with *Hin*dIII and *Xho*I, followed by incorporation into pE, yielding pBLT1E. pBLT2E was constructed according to the construction method of pBLT1E.

Transformation of *B. licheniformis*

A method of electrotransformation was used to transform plasmids into *B. licheniformis*. A total of 8 μg-10 μg of plasmid was added to the *B. licheniformis* competent cells and mixed. Then, the mixed competent cells were added to a precooled 0.1 cm Gene Pulser cuvette and placed on ice for 5 min. Next, the 0.1 cm Gene Pulser cuvette was placed into an electroporation apparatus and shocked with 2,100 V. After the electric shock was administered, 900 μL of recovery medium (LB + 0.5 M sorbitol + 0.38 M mannitol) was immediately added, and the cells were cultured at 37°C and 100 rpm for 3 h before being applied to the corresponding antibiotic plates.

Fluorescence Measurement of eGFP

The BlspTE strain was cultured overnight at 37°C and 250 rpm, and 3% culture was inoculated into the fermentation broth. After 8 h of growth, 15 g/L trehalose was added for inducible expression. The sample was measured for OD600 and fluorescence intensity after 16 h of adding trehalose. Then, 100 μL of fermentation broth was centrifuged to obtain the cell pellet, then the cell pellet was rinsed twice with phosphate buffered saline (PBS) solution at pH 7.4, and the final OD600 was diluted to 0.5. Next, 200 μL of diluted suspension was added to the 96-well microtiter plate (Corning). The 96-well microtiter plate was placed in a TECAN-SparK plate reader (Tecan, Männedorf, Switzerland), which calculated the final value using an absorption wavelength of 485 nm, excitation wavelength of 535 nm, and gain value of 100. The formula FI (AU/OD) = 2×(FVt-FVr), where FVt refers to the fluorescence value measured by the target strain, and FVr represents the fluorescence value measured by the control strain BlsPE, was used to evaluate the fluorescence intensity.

Electrophoretic Mobility Shift Assays

DNA probes were amplified using high-fidelity DNA polymerase (Vazyme Biotech Co., Ltd. 2 × Phanta Master Mix) and primers were labeled with biotin. Biotin-labeled probes were purified by agarose gel electrophoresis. 10 nM biotin-labeled probes were incubated with different concentrations of CcpA in binding buffer (10 mM Tris-HCl (pH 7.4), 1 mM DTT, 1 mM EDTA, 50 mM KCl, 0.05 μg/μL poly (dI-dC), 1 mM $MgCl_2$), and the reaction solution was put through a full reaction at 25°C for 20 min. After the reaction was completed, the samples were separated by electrophoresis using 4% acrylamide gels in 0.5 × Tris-borate EDTA (TBE) buffer. The samples were electroblotted from the acrylamide gels onto nylon membrane (Beyotime, FFN15), and immobilized by UV crosslinking. The nylon membrane was washed and detected using Chemiluminescent EMSA Kit (Beyotime, GS009) according to manufacturing protocol. Gel imaging and analysis were performed using ChemiDoc XRS (Bio-Rad, U.S.A).

Statistical analysis

The sample size was $n \geq 3$ for biology experiments. A student's tests (*$P \leq 0.05$; **$P \leq 0.01$, ***$P \leq 0.001$) were performed for statistical analysis.

Fig. S1

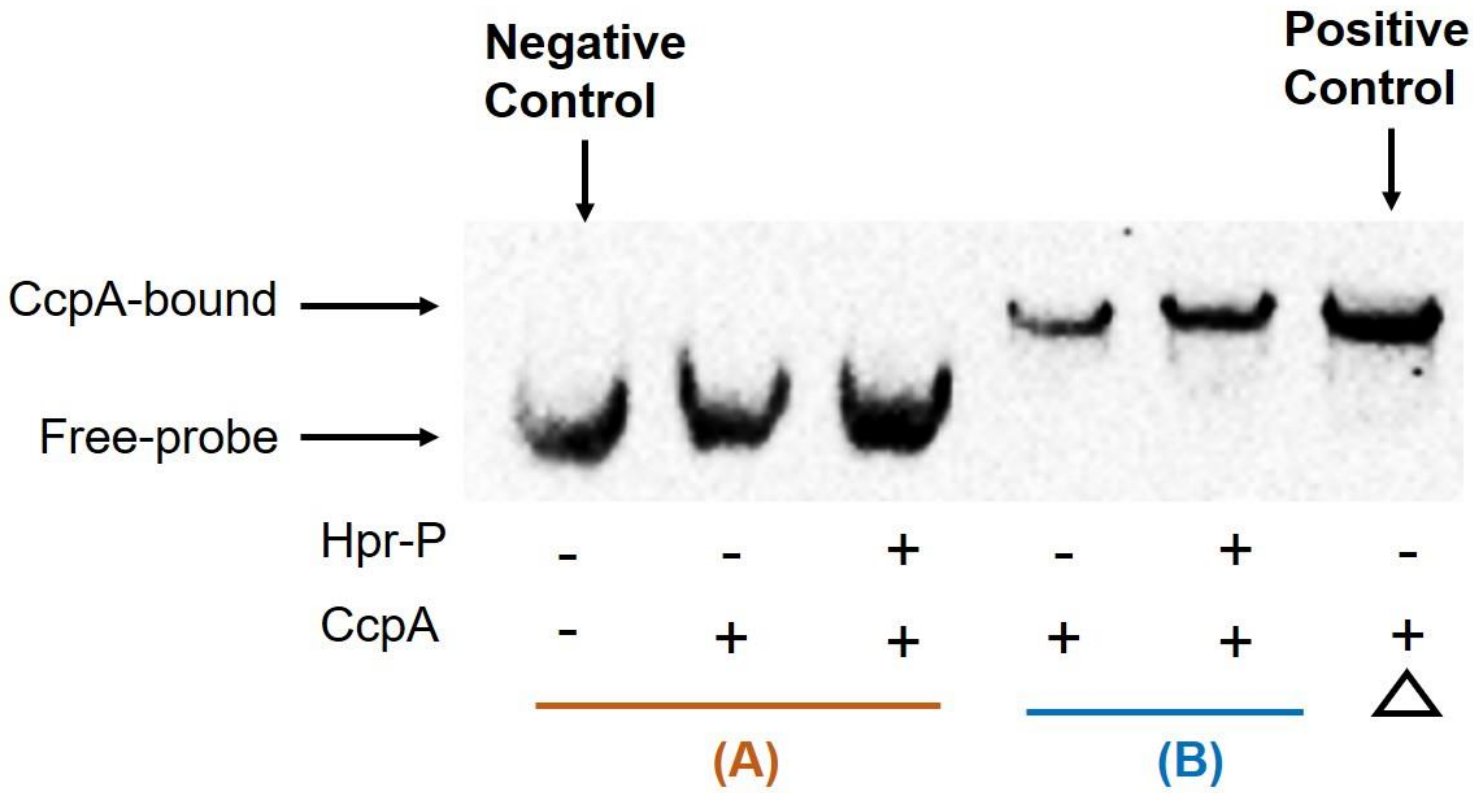


Figure S1: EMSA of CcpA or CcpA-Hpr-P protein for fragment A and B. Related to Figure 1.

The *hpr* (c16360) gene of *Bacillus licheniformis* CICIM B1391 was cloned and then inserted into pET28a vector (between *Nde*I and *Eco*RI), to obtain the gene expression vector of pET28a-*hpr*. The recombinant strain *E.coli* BL21 (DE3) containing expression vector was grown overnight in 15 mL LB-Kanamycin at 37℃, 200 rpm. A 3 % of the culture was inoculated into Terrific Broth (TB) at 37℃ until OD600 of 0.4-0.5. The Hpr protein was induced by 0.1mM IPTG for 10 h at 25℃. The cells were collected by centrifugation at 8,000 × g for 5 min. The Hpr protein was purified by the Kit (Mag-Beads His-Tag Protein Purification, Sangon Biotech, C650033) according to the manufacture's protocols. The purify of the Hpr Protein was analyzed through Tris-Tricine SDS PAGE.

Hpr protein phosphorylation was performed in 20 mM Tris-Cl buffer (pH 7.0) supplemented with 1 mM $MgCl_2$, 2 mM NaCl and 5 mM ATP. 27 μg Hpr protein and 0.15 μg HprK protein was added to the reaction system at 37℃ for 10 min. The reaction was ended at 75℃ for 5 min. 10% phos-tag SDS PAGE (Boppard, 193-16711) was used to determine whether Hpr phosphorylation.

The EMSA Lane (left→right): 10 nM fragment A; 1.2μM CcpA+10 nM fragment A; 1.2 μM CcpA +1.2 μM Hpr-P+10 nM fragment A; 1.2μM CcpA+10 nM fragment B; 1.2 μM CcpA +1.2 μM Hpr-P+10 nM fragment B; 1.2μM CcpA + 10 nM Positive control probe.

Fig. S2

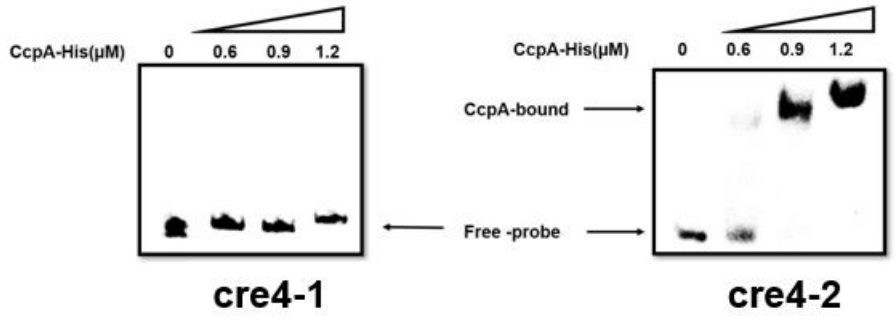


Figure S2: EMSA of CcpA protein binding to two fragments (E6, E7) labelled with 5'-biotin. Related to Figure 1.

Fragment E6 and Fragment E7, containing cre4-1 (AGCGTT-aaggaactttcaga-AACGCT) and cre4-2 (AGCTTT-aaggaactttcaga-AAAGCT). Increasing concentrations of CcpA (0 μM, 0.6 μM, 0.9 μM, 1.2 μM) were incubated with 10 nM E6 or E7 before the reaction run on EMSA gel.

Fig. S3

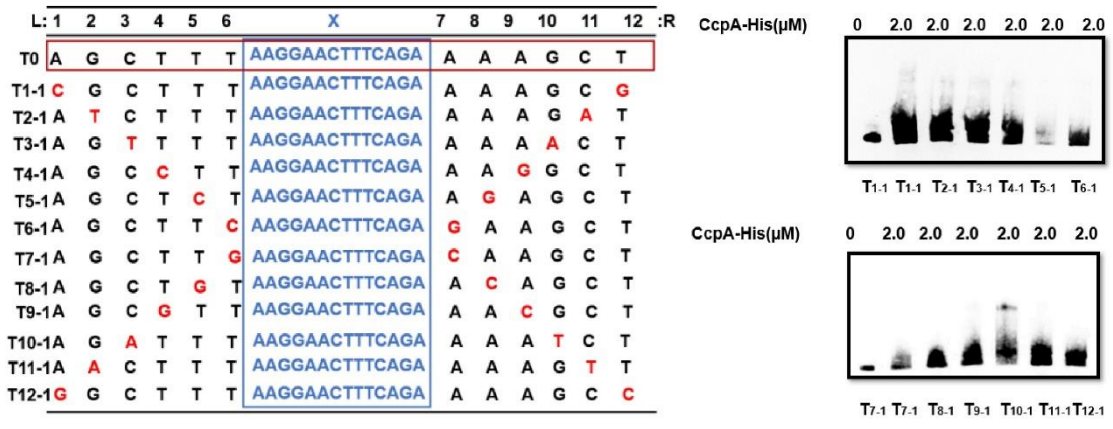


Figure S3: Nucleotide mutations within the 12-bp symmetrical regions of CRE_{Tre}. Related to Figure 3.

Two symmetrical bases in symmetrical regions were mutated based on fragment H1, resulting 12 derivative probes (T1-1, T2-1, T3-1, T4-1, T5-1, T6-1, T7-1, T8-1, T9-1, T10-1, T11-1, and T12-1). 2.0 μM CcpA was incubated with 10 nM T1-1, T2-1, T3-1, T4-1, T5-1, T6-1, T7-1, T8-1, T9-1, T10-1, T11-1, and T12-1before the reaction run on EMSA gel.

Fig. S4

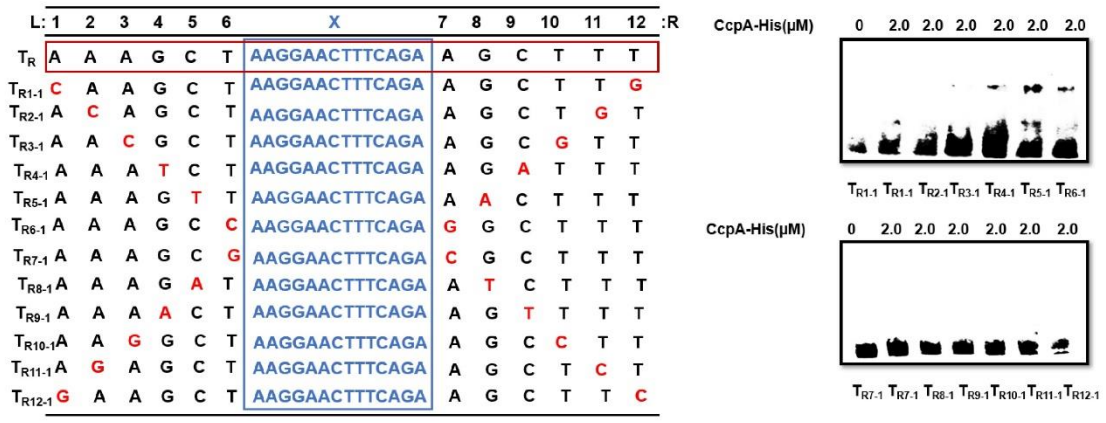


Figure S4: Further point mutation of the 12-bp symmetrical region of $CRE_{Tre(R)}$. Related to Figure 6.

Two symmetrical bases in symmetrical regions were mutated based on fragment H1, resulting 12 derivative probes (TR1-1, TR2-1, TR3-1, TR4-1, TR5-1, TR6-1, TR7-1, TR8-1, TR9-1, TR10-1, TR11-1, and TR12-1). 2.0 μM CcpA was incubated with 10 nM (TR1-1, TR2-1, TR3-1, TR4-1, TR5-1, TR6-1, TR7-1, TR8-1, TR9-1, TR10-1, TR11-1, and TR12-1), and T12-1before the reaction run on EMSA gel.

Fig. S5

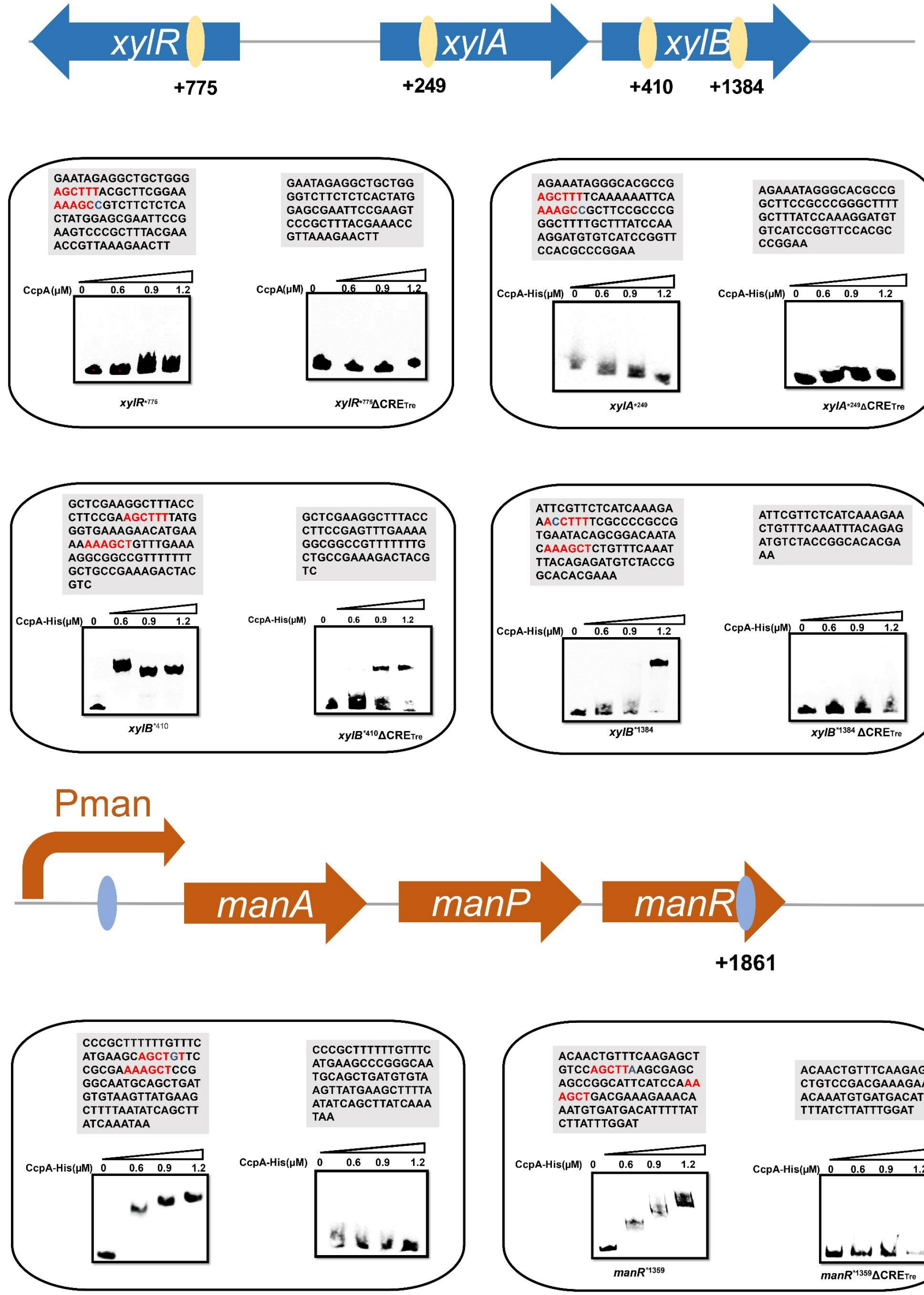


Figure S5: Potential CRE_{Tre} sites in xylose operon and mannose operon. Related to Figure 8.

The putative CRE_{Tre} was annotated with red. Increasing concentrations of CcpA (0 μM, 0.6 μM, 0.9 μM, 1.2 μM) were incubated with 10 nM fragments before the reaction run on EMSA gel.

Table S1
Bacterial Strains and Plasmids Used in This Study. Related to Figure 4.

Strain or plasmid	Description	Reference
Strains		
*Escherichia coli*JM109	F′, traD36, proAB + lacIq, Δ(lacZ), M15/Δ (lac-proAB), gln V44, e14−, gyrA96, recA1, relA1, endA1, thi, hsdR17 (CICIM B0012)	CICIM-CU
Bacillus licheniformis CICIM B1391	wild-type (CICIM B1391)	CICIM-CU
Bacillus licheniformis CA	*B. licheniformis* CICIM B1391, Δ*ccpA*	Laboratory construct
BlspHY	*B. licheniformis* CICIM B1391, harboring pHY300-PLK	this work
BlspE	*B. licheniformis* CICIM B1391, harboring pE	this work
BlspTE	*B. licheniformis* CICIM B1391, harboring pBLTE	this work
BlspT1E	*B. licheniformis* CICIM B1391, harboring pBLT1E	this work
BlspT2E	*B. licheniformis* CICIM B1391, harboring pBLT2E	this work
BlspDE	*B. licheniformis* CICIM B1391, harboring pBLDE	this work
BlspAE	*B. licheniformis* CICIM B1391, harboring pBLAE	this work
BlspHY1	*B. licheniformis* CA, harboring pHY300-PLK	this work
BlspE1	*B. licheniformis* CA, harboring pE	this work
BlspTE1	*B. licheniformis* CA, harboring pBLTE	this work
BlspT1E1	*B. licheniformis* CA, harboring pBLT1E	this work
BlspT2E1	*B. licheniformis* CA, harboring pBLT2E	this work
BlspDE1	*B. licheniformis* CA, harboring pBLDE	this work
BlspAE1	*B. licheniformis* CA, harboring pBLAE	this work
Plasmids		this work
pMD18-T-simple	*E. coli* cloning vector, Ap^R	TaKaRa
pHY300-PLK	*E. coli*/*Bacillus* shuttle vector, Ap^R/Tet^R	TaKaRa
pE	pHY300-PLK derivative with egfp	this work
pBLTE	pE derivative with P*treA* from *B. licheniformis* CICIM B1391	this work
pBLT1E	pE derivative with P*treA,* in which CRE site 'TGAAAGCGCTATAA' was changed to 'AGCTTT-AT-AAAGCT'	this work
pBLT2E	pE derivative with P*treA,* in which CRE site 'TGAAAGCGCTATAA' was changed to 'AAAGCT-AT-AGCTTT'	this work
pBLDE	pE derivative with P*dnaP* from *B. licheniformis* CICIM B1391	this work
pBLAE	pE derivative with P5A2 from *B. licheniformis* CICIM B1391	this work

Ap^R, ampicillin resistance; Tet^R, tetracycline resistance; Kan^R, kanamycin resistance
CICIM-CU: Culture and Information Center of Industrial Microorganisms of China Universities.

Table S2
Primers Used to Construct Recombinant Plasmids. Related to Figure 4.

Primers	Sequence (5'-3')	Restriction site
eGFPF	CCGCTCGAGatgggtcgcggatccatg	*Xho*I
eGFPR	TCCCCCGGGtcacacgtggtggtggtg	*Sma*I
P*treA*-F	CCCAAGCTTatctcagccggttgttcc	*Hin*dIII
P*treA*-R	CCGCTCGAGttccaatccctccttctc	*Xho*I
P*treA*-CRETre-F	AGCTTTATAAAGCTaaatatgttgactacttgt	
P*treA*-CRETre-R	AGCTTTATAAAGCTaaaataaaaaaagcccgg	
P*treA*-CRETre(R)-F	AAAGCTATAGCTTTaaatatgttgactacttgt	
P*treA*-CRETre(R)-R	AAAGCTATAGCTTTaaaataaaaaaagcccgg	
P*danP*-F	CCCAAGCTTacgcttagggtagcgaatcga	*Hin*dIII
P*dnaP*-R	CCGCTCGAGcggcctgaacactctatata	*Xho*I
P5A2-F	CCCAAGCTTagggaggggaaagctaaat	*Hin*dIII
P5A2-R	CCGCTCGAGctaggattttacctccctttc	*Xho*I
treR-199-F	atgaagatcaacaagtatat	
treR-169-F	aaacaaattgcagaacaaat	
treR-139-F	atactgaatgccggagata	
treR-109-F	gaaaacgaccttgccgaac	
treR-79-F	tcgagggaaacggttcgga	
treR-49-F	gtgcttgcccagaacggat	
treR-199-R(5'-biotin)	ctgacccttttcctttgatct	
treR-202-F	agatcaaaggaaaagggt	
treR-172-F	tcgtgctgcacagggaga	
treR-142-F	cggtttccggtttggtcag	
treR-112-F	tttcagaaacgcttggcaa	
treR-82-F	cgacggtccacgagttcgg	
treR-52-F	cggatgcatatattcaaaa	
treR-202-R(5'-biotin)	acgacccgccatacctctt	
treR-196-F	ttctggagatgaagaggtatg	
treR-196-R(5'-biotin)	gaccacaatttccttatg	
treR-194-F	atgcccataaggaaattg	
treR-194-R(5'-biotin)	ctacttttctcttcttgca	
P*treA*-88-F	atctcagccggttgttcccggc	
P*treA*-88-R(5'-biotin)	aaataaaaaaagcccggccg	
P*treA*-102-F	ttgaaagcgctataaaaat	
P*treA*-102-R(5'-biotin)	ttccaatccctccttctc	
PtreA-102ΔCRE-F	taaatatgttgactacttg	
tre-26-1-F	AGCTTTaaggaactttcagaAAAGCTatctcagccggttgttcccg	
tre-20-2-F	AGCTTTaaggttcagaAAAGCTatctcagccggttgttcccg	
tre-12-3-F	AGCTTTAAAGCTatctcagccggttgttcccg	
tre-26-4-F	CCCGGGaaggaactttcagaAAAGCTatctcagccggttgttcccg	
tre-26-5-F	AGCTTTaaggaactttcagaCCCGGGatctcagccggttgttcccg	
tre-26-6-F	CCCGGGaaggaactttcagaCCCGGGatctcagccggttgttcccg	
tre(R)-26-1-F	AAAGCTaaggaactttcagaAGCTTTatctcagccggttgttcccg	
tre(R)-20-2-F	AAAGCTaaggttcagaAGCTTTatctcagccggttgttcccg	
tre(R)-12-3-F	AAAGCTAGCTTTatctcagccggttgttcccg	
tre(R)-26-4-F	CCCGGGaaggaactttcagaAGCTTTatctcagccggttgttcccg	
tre(R)-26-5-F	AAAGCTaaggaactttcagaCCCGGGatctcagccggttgttcccg	
tre(R)-26-6-F	CCCGGGaaggaactttcagaCCCGGGatctcagccggttgttcccg	
T1-F	CGCTTTaaggaactttcagaAAAGCTatctcagccggttgttcccg	
T2-F	ATCTTTaaggaactttcagaAAAGCTatctcagccggttgttcccg	
T3-F	AGTTTTaaggaactttcagaAAAGCTatctcagccggttgttcccg	
T4-F	AGCCTTaaggaactttcagaAAAGCTatctcagccggttgttcccg	
T5-F	AGCTCTaaggaactttcagaAAAGCTatctcagccggttgttcccg	
T6-F	AGCTTCaaggaactttcagaAAAGCTatctcagccggttgttcccg	
T7-F	AGCTTTaaggaactttcagaCAAGCTatctcagccggttgttcccg	
T8-F	AGCTTTaaggaactttcagaACAGCTatctcagccggttgttcccg	
T9-F	AGCTTTaaggaactttcagaAACGCTatctcagccggttgttcccg	
T10-F	AGCTTTaaggaactttcagaAAATCTatctcagccggttgttcccg	
T11-F	AGCTTTaaggaactttcagaAAAGTTatctcagccggttgttcccg	
T12-F	AGCTTTaaggaactttcagaAAAGCCatctcagccggttgttcccg	
TR-1-F	CAAGCTaaggaactttcagaAGCTTTatctcagccggttgttcccg	
TR-2-F	ACAGCTaaggaactttcagaAGCTTTatctcagccggttgttcccg	
TR-3-F	AACGCTaaggaactttcagaAGCTTTatctcagccggttgttcccg	
TR-4-F	AAATCTaaggaactttcagaAGCTTTatctcagccggttgttcccg	

TR-5-F	AAAGTTaaggaactttcagaAGCTTTatctcagccggttgttcccg
TR-6-F	AAAGCCaaggaactttcagaAGCTTTatctcagccggttgttcccg
TR-7-F	AAAGCTaaggaactttcagaCGCTTTatctcagccggttgttcccg
TR-8-F	AAAGCTaaggaactttcagaATCTTTatctcagccggttgttcccg
TR-9-F	AAAGCTaaggaactttcagaAGTTTTatctcagccggttgttcccg
TR-10-F	AAAGCTaaggaactttcagaAGCCTTatctcagccggttgttcccg
TR-11-F	AAAGCTaaggaactttcagaAGCTCTatctcagccggttgttcccg
TR-12-F	AAAGCTaaggaactttcagaAGCTTCatctcagccggttgttcccg
T1-1-F	CGCTTTaaggaactttcagaAAAGCGatctcagccggttgttcccg
T2-1-F	ATCTTTaaggaactttcagaAAAGATatctcagccggttgttcccg
T3-1-F	AGTTTTaaggaactttcagaAAAACTatctcagccggttgttcccg
T4-1-F	AGCCTTaaggaactttcagaAAGGCTatctcagccggttgttcccg
T5-1-F	AGCTCTaaggaactttcagaAGAGCTatctcagccggttgttcccg
T6-1-F	AGCTTCaaggaactttcagaGAAGCTatctcagccggttgttcccg
T7-1-F	AGCTTGaaggaactttcagaCAAGCTatctcagccggttgttcccg
T8-1-F	AGCTGTaaggaactttcagaACAGCTatctcagccggttgttcccg
T9-1-F	AGCGTTaaggaactttcagaAACGCTatctcagccggttgttcccg
T10-1-F	AGATTTaaggaactttcagaAAATCTatctcagccggttgttcccg
T11-1-F	AACTTTaaggaactttcagaAAAGTTatctcagccggttgttcccg
T12-1-F	GGCTTTaaggaactttcagaAAAGCCatctcagccggttgttcccg
TR-1-1-F	CAAGCTaaggaactttcagaAGCTTGatctcagccggttgttcccg
TR-2-1-F	ACAGCTaaggaactttcagaAGCTGTatctcagccggttgttcccg
TR-3-1-F	AACGCTaaggaactttcagaAGCGTTatctcagccggttgttcccg
TR-4-1-F	AAATCTaaggaactttcagaAGATTTatctcagccggttgttcccg
TR-5-1-F	AAAGTTaaggaactttcagaAACTTTatctcagccggttgttcccg
TR-6-1-F	AAAGCCaaggaactttcagaGGCTTTatctcagccggttgttcccg
TR-7-1-F	AAAGCGaaggaactttcagaCGCTTTatctcagccggttgttcccg
TR-8-1-F	AAAGATaaggaactttcagaATCTTTatctcagccggttgttcccg
TR-9-1-F	AAAACTaaggaactttcagaAGTTTTatctcagccggttgttcccg
TR-10-1-F	AAGGCTaaggaactttcagaAGCCTTatctcagccggttgttcccg
TR-11-1-F	AGAGCTaaggaactttcagaAGCTCTatctcagccggttgttcccg
TR-12-1-F	GAAGCTaaggaactttcagaAGCTTCatctcagccggttgttcccg
xylB-410-F	agcttttatgggaaagaaca
xylB-410ΔCRETre-F	gtttgaaaaggcggccgt
xylB-410-R(5'-biotin)	attcgctgtgaagacgccc
xylB-1384-F	acctttcgccccgccgtga
xylB-1384ΔCRETre-F	ctgtttcaaatttacagaga
xylB-1384-R(5'-biotin)	ttattcccgaaagctcgcca
xylR-775-F	agctttacgcttcggaaaaagc
xylR-775ΔCRETre-F	gtcttctctcactatggag
xylR-775-R(5'-biotin)	gaaaccgttaaagaacttgc
xylA-249-F	agcttttcaaaaaattca
xylA-249ΔCRETre-F	gcttccgcccgggctttt
xylA-249-R(5'-biotin)	atgacacatcctttggataaagc
levR-1359-F	agctttttaatcagaaaat
levR-1359ΔCRETre-F	tcaaaaccacagcttgctca
levR-1359-R(5'-biotin)	aattaaaggaagtcgctgaa
manR-1359-F	agcttttggatgaatgccgg
manR-1359ΔCRETre-F	ggacagctcttgaaacagt
manR-1359-R(5'-biotin)	ctttatgctggctgtcaaac
Ptre--48-F	cgctttcaaaaataaaaaaa
Ptre--71-F	cggccgttcccttcataa
Ptre--48-R(5'-biotin)	atctcagccggttgttcccg